

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M nucleic - nucleic search, using sw model

run on: October 22, 2003, 07:27:11 ; Search time 5513 Seconds
(without alignments)
11197.647 Million cell updates/sec

title: US-10-006-852-1
effect score: 1509
sequence: 1 atgggtctctcccacgcggt.....agaagagtgtatctgtctaa 1509

scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

searched: 2888711 seqs, 20454813386 residues

total number of hits satisfying chosen parameters: 5777422

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_stg.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_pa.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_ov.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1509	100.0	1509	8	ATU10034	U10034 Arabidopsis
2	1507.4	99.9	1509	6	AX507115	AX507115 Sequence
3	1507.4	99.9	1509	6	AX651552	AX651552 Sequence
4	1507.4	99.9	1509	8	BT001047	BT001047 Arabidops
5	1507.4	99.9	1813	8	AF094464	AF094464 Arabidops
6	907.4	60.1	1776	8	AF353615	AF353615 Nicotiana
7	883.4	58.5	1705	8	AF020425	AF020425 Nicotiana
8	881.8	58.4	1785	8	PETGADX	L16797 Petunia hyb
9	881.8	58.4	1785	8	PETGUDRCA	L16977 Petunia hyb
10	881.8	58.4	1729	8	NTU54774	U54774 Nicotiana t
11	877	58.1	1672	8	AF352732	AF352732 Nicotiana
12	870.2	57.7	1485	8	AY124873	AY124873 Arabidops
13	870.2	57.7	1676	8	AY146665	AY146665 Arabidops
14	870.2	57.7	1681	8	ATU49937	U49937 Arabidops
15	870.2	57.7	1695	8	AF428372	AF428372 Arabidops
16	868.6	57.5	1685	8	AF428294	AF428294 Arabidops
17	864.2	57.3	1745	8	AF506366	AF506366 Nicotiana
18	861.4	57.1	1650	8	AY081259	AY081259 Arabidops
19	857.8	56.8	1771	8	AF020424	AF020424 Nicotiana
20	854.2	56.6	1482	6	AX412801	AX412801 Sequence
21	854.2	56.6	1482	6	AX506242	AX506242 Sequence
22	854.2	56.6	1482	6	AX651853	AX651853 Sequence
23	854.2	56.6	1482	8	AY124860	AY124860 Arabidops
24	854.2	56.6	1792	8	AF361836	AF361836 Arabidops
25	838.2	55.5	1712	8	AK118125	AK118125 Arabidops
26	834.2	55.3	1479	8	AX654603	AX654603 Sequence
27	834.2	55.3	1502	8	AY187941	AY187941 Oryza sat
28	770	51.0	1605	8	AY084689	AY084689 Arabidops
29	769	51.0	1779	8	AB056060	AB056060 Oryza sat
30	736.6	48.8	1783	6	AB39864	AB39864 Sequence 8
31	736.6	48.8	1783	6	AR070945	AR070945 Sequence
32	736.6	48.8	1783	8	LEGDL	X80840 L.esculentu
33	711.4	47.1	1180	6	BD017613	BD017613 Method fo
34	709.6	47.0	1455	6	AX653206	AX653206 Sequence
35	670.6	44.4	1781	8	AB056061	AB056061 Oryza sat
36	560.2	37.1	1467	6	AX653024	AX653024 Sequence
37	487.6	32.3	75188	8	AB005238	AB005238 Arabidops
38	432.4	28.7	136098	8	AC137696	AC137696 Genomic s
39	415.2	27.5	3983	8	AB056063	AB056063 Oryza sat
40	415.2	27.5	147227	8	OSJN00013	AL606448 Oryza sat
41	393	26.0	5160	6	AX461514	AX461514 Sequence
42	393	26.0	81490	8	FL2P19	AC009513 Sequence
43	386.2	25.6	1545	6	BD008061	BD008061 Glutamate
44	385.8	25.6	76976	8	AC006532	AC006532 Arabidops
45	380.4	25.2	906	8	AF358767	AF358767 Oryza sat

ALIGNMENTS

RESULT 1
ATU10034
LOCUS ATU10034 1509 bp mRNA linear PLN 30-MAR-1996
DEFINITION Arabidopsis thaliana Columbia glutamate decarboxylase (GAD) mRNA,
complete cds.
ACCESSION U10034
VERSION U10034.1 GI:497978
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicot;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1509)
AUTHORS Arazi,T., Baum,G., Snedden,W.A., Shelp,B.J. and Fromm,H.

TITLE Molecular and biochemical analysis of calmodulin interactions with the calmodulin-binding domain of plant glutamate decarboxylase
 JOURNAL Plant Physiol. 108 (2), 551-561 (1995)
 MEDLINE 95334488
 PUBMED 7610159
 PREFERENCE 2 (bases 1 to 1509)
 AUTHORS Fromm, H.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAY-1994) Hillel Fromm, Department of Plant Genetics, Weizmann Institute of Science, Rehovot 76100, Israel
 FEATURES Location/Qualifiers
 source 1..1509
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strains="Columbia"
 /db_xref="taxon:3702"
 /clone="35B3T7"
 1..1509
 /gene="GAD"
 1..1509
 /gene="GAD"
 /EC_number="4.1.1.15"
 /note="GAD1 isoform; similar to the GAD2 isoform of Arabidopsis thaliana, GenBank Accession Number U49937"
 /codon_start=1
 /product="glutamate decarboxylase"
 /protein_id="AAA93132.1"
 /db_xref="gi:497979"
 /translation="MWLSHAVSESDVSHVSTFASRYVRTSLPRFMPENSIPKEAAVQ IINDELMLDGNPRLNLASFTVWVDECDKLMSINKVYVDMDEYPTVELQNRVY MIAHFNAPLEAEATAVGVGVGSSSEALMLAGLAFKRWKQKRAEKQKPVDKNIWV ANVQVWEKPARVEVELKEVLSGYYVMDPQAVDVBENTICVADILGSLNGBF EDVLLDNLLEKKNKGTWDTPHVDAAASGDFIAPFLYPELWDFRPLVKSINVSFH KYGLVAGIGWVIWRNKEDLPELIFHNYLGDAPFTLNFPSKGSQVIAQYYLLIR LGHEGYRNMENRENMLVRELEKTERFNIVSKDEGVPLVAFSLKSDSCHTEFELS DMLRRYGVIQVAYTWPNAOHIIVLRVIREDFRSLAERLVIDIEKVMRELDLPSR VHKISLGOEKSESNDLNVTKKSDIDKQDIIITGWKQVADRKTKTSGIC"
 misc_feature 1417..1506
 /gene="GAD"
 /note="calmodulin-binding domain"
 433 a 321 c 385 g 370 t

Query Match 100.0%; Score 1509; DB 8; Length 1509;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTTGCTCTCCACGGGTATCGGAGTGGAGTCCGTCCTCCACTCCACATTCGCATCA 60
 |||||
 1 ATGTTGCTCTCCACGGGTATCGGAGTGGAGTCCGTCCTCCACTCCACATTCGCATCA 60
 |||||
 61 CGTTACGTCGTTACTCTTCTAGTTCAAGATGCGGAAAACCTCGATTCTTAAGAA 120
 |||||
 61 CGTTACGTCGTTACTCTTCTAGTTCAAGATGCGGAAAACCTCGATTCTTAAGAA 120
 |||||
 121 GCGGGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTCTCACTTA 180
 |||||
 121 GCGGGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTCTCACTTA 180
 |||||
 181 GCCTCTTTGTGACGACATGATGAGGCTGAGTGTGATATAACTCATCATCTCTCCATC 240
 |||||
 181 GCCTCTTTGTGACGACATGATGAGGCTGAGTGTGATATAACTCATCATCTCTCCATC 240
 |||||
 241 AACAAACTATGTTGACATGACGAGTACCCCGTCAACCCGAACTTCAGAACCGATGT 300
 |||||
 241 AACAAACTATGTTGACATGACGAGTACCCCGTCAACCCGAACTTCAGAACCGATGT 300
 |||||
 301 GTGAACATGATGACATCTAATCAACCGTTAGAGGGGGAGACCCCGTCCGA 360
 |||||
 301 GTGAACATGATGACATCTAATCAACCGTTAGAGGGGGAGACCCCGTCCGA 360
 |||||
 361 GTAGGAACCGTTGATCATCGGAGGCCATAATGTTGGCCGTTTGGCCCTTCAAGGTAAA 420
 |||||

361 GTAGGAACCGTTGATCATCGGAGGCCATAATGTTGGCCGTTTGGCCCTTCAAGCGTAAA 420
 |||||
 421 TGGCAGAAACAAAGCGCAAAAGCTGAAGGCAAAACCGCTCGATAAAAACCAACATTTGTCACCCGGA 480
 |||||
 421 TGGCAGAAACAAAGCGCAAAAGCTGAAGGCAAAACCGCTCGATAAAAACCAACATTTGTCACCCGGA 480
 |||||
 481 GCCAAATGTTCAAGTGTGTTGGGAAAATTCGTTAGTACTCTTTCAGGTTTGAACCTTAAGAA 540
 |||||
 481 GCCAAATGTTCAAGTGTGTTGGGAAAATTCGTTAGTACTCTTTCAGGTTTGAACCTTAAGAA 540
 |||||
 541 GTGAAATGAGTGAAGGATATGATGAGGACCCCTCAACAGCTGTTGATATGGTTGAT 600
 |||||
 541 GTGAAATGAGTGAAGGATATGATGAGGACCCCTCAACAGCTGTTGATATGGTTGAT 600
 |||||
 601 GAGAACACCAATTTGTTGGGACATTTCTTTGGTTCCACTCTTAATGAGAAATTCGAAGAT 660
 |||||
 601 GAGAACACCAATTTGTTGGGACATTTCTTTGGTTCCACTCTTAATGAGAAATTCGAAGAT 660
 |||||
 661 GTTAAACTCTTTGAAACGATCTCTTTGGTCGAAAAGAACAAAACCGGATGGGATACACCA 720
 |||||
 661 GTTAAACTCTTTGAAACGATCTCTTTGGTCGAAAAGAACAAAACCGGATGGGATACACCA 720
 |||||
 721 ATCCACGTGATCGGCAAGTGGAGGATTCATTGCAACCGTTTTTGTATCCGGAAATGGAA 780
 |||||
 721 ATCCACGTGATCGGCAAGTGGAGGATTCATTGCAACCGTTTTTGTATCCGGAAATGGAA 780
 |||||
 781 TGGACTTTAGACTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCAAAAGTATGGACTT 840
 |||||
 781 TGGACTTTAGACTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCAAAAGTATGGACTT 840
 |||||
 841 GTCTACCGCAGGGATTTGGTTGGTGAATCTGGAGAAAACAAGAGGATTTGCCTGAGGAATC 900
 |||||
 841 GTCTACCGCAGGGATTTGGTTGGTGAATCTGGAGAAAACAAGAGGATTTGCCTGAGGAATC 900
 |||||
 901 ATCTTCCATATCAATATCTTTGGTGTGACCAACCCACTTTACTCTCAATTTCTCCAAA 960
 |||||
 901 ATCTTCCATATCAATATCTTTGGTGTGACCAACCCACTTTACTCTCAATTTCTCCAAA 960
 |||||
 961 GFTTCAAGTCAAGTCAATGCTCAATACTACCAACTTATCCGATTTGGGCCACGAGGTTAC 1020
 |||||
 961 GFTTCAAGTCAAGTCAATGCTCAATACTACCAACTTATCCGATTTGGGCCACGAGGTTAC 1020
 |||||
 1021 AGAAATGTGATGAGAAATTCAGAGAAATGATCGTCTTAAGGAAAGACTTTGAGAAG 1080
 |||||
 1021 AGAAATGTGATGAGAAATTCAGAGAAATGATCGTCTTAAGGAAAGACTTTGAGAAG 1080
 |||||
 1081 ACAGAAAGTTCACACATCGTCTCAAAAGCAGGAGTGCCTTGTGCGTTTCTCCTTG 1140
 |||||
 1081 ACAGAAAGTTCACACATCGTCTCAAAAGCAGGAGTGCCTTGTGCGTTTCTCCTTG 1140
 |||||
 1141 AAAGATAGCAGTGTCACTGAGTTCGAAAATCTCCGACATGCTTCGAGGATGATGATGG 1200
 |||||
 1141 AAAGATAGCAGTGTCACTGAGTTCGAAAATCTCCGACATGCTTCGAGGATGATGATGG 1200
 |||||
 1201 ATAGTCCGCGCTTACACATGCTCCAAAATGCAACACATCACTGTTCTTCTCGTGGTT 1260
 |||||
 1201 ATAGTCCGCGCTTACACATGCTCCAAAATGCAACACATCACTGTTCTTCTCGTGGTT 1260
 |||||
 1261 ATCAGAGAAATTTCTCGAAACACTCGTGTGAGACTTGTGATCGATATAGAAAAGT 1320
 |||||
 1261 ATCAGAGAAATTTCTCGAAACACTCGTGTGAGACTTGTGATCGATATAGAAAAGT 1320
 |||||
 1321 ATCGGTGAGTCTGATGAGTTCCTTTCGAGAGTATCAAAAATATCACTTTGACAAAG 1380
 |||||
 1321 ATCGGTGAGTCTGATGAGTTCCTTTCGAGAGTATCAAAAATATCACTTTGACAAAG 1380
 |||||
 1381 AAGAGTCAATTAACAGCGATTAACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1440
 |||||
 1381 AAGAGTCAATTAACAGCGATTAACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1440
 |||||
 1441 CAGAGAGATATCATCACTGGCTGAAAGATTTGTTGCGCGACAGGAAAGAGAGAGTGGT 1500
 |||||
 1441 CAGAGAGATATCATCACTGGCTGAAAGATTTGTTGCGCGACAGGAAAGAGAGAGTGGT 1500
 |||||

CESSION AX651552
 RSION AX651552.1 GI:29154370
 WORDS Arabidopsis thaliana (thale cress)
 URCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 I
 FERENCE Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
 AUTHORS Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
 TITLE Plant genes involved in defense against pathogens
 JOURNAL Patent: WO 03000898-A 360 03-JAN-2003;
 Syngenta Participations AG (CH)
 ATURES Location/Qualifiers
 source 1. 1509
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 SE COUNT 432 a 322 c 385 g 370 t
 IGIN

Query Match 99.9%; Score 1507.4; DB 6; Length 1509;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGTCTCTCCACCGCGTATCGGAGTCCGACGCTCCGTCCTCCACTCCACATTCGCATCA 60
 1 ATGGTCTCTCCACCGCGTATCGGAGTCCGACGCTCCGTCCTCCACTCCACATTCGCATCA 60

61 CGTTACGTCCTTCACTTCTTAGGTTCAAGATGCCGGAAACTCGATTCTTAAGGAA 120
 61 CGTTACGTCCTTCACTTCTTAGGTTCAAGATGCCGGAAACTCGATTCTTAAGGAA 120

121 CCGCGTATCAGATCAACACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 121 CCGCGTATCAGATCAACACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180

181 GCCTCCTTTGACGACATGATGAGTCCGATGATGATGATGATGATGATGATGATGATG 240
 181 GCCTCCTTTGACGACATGATGAGTCCGATGATGATGATGATGATGATGATGATGATG 240

241 AACAAAGACTATGTTGACATCGACGAGTACCCGTCACCCGGAATTCAGAACCGATGT 300
 241 AACAAAGACTATGTTGACATCGACGAGTACCCGTCACCCGGAATTCAGAACCGATGT 300

301 GTGAACATGATGACATCTATTCAATGCAATGATGATGATGATGATGATGATGATGATG 360
 301 GTGAACATGATGACATCTATTCAATGCAATGATGATGATGATGATGATGATGATGATG 360

361 GTAGGAACCGTTGATGATCGGAGCCATATGTTGGCCGTTTGGCCCTTCAAGCGTAAA 420
 361 GTAGGAACCGTTGATGATCGGAGCCATATGTTGGCCGTTTGGCCCTTCAAGCGTAAA 420

421 TGGCAGAACGCGCAAGCTGAAGCAAAACCCGTCGATAAACCCCAACTGTCACCGGA 480
 421 TGGCAGAACGCGCAAGCTGAAGCAAAACCCGTCGATAAACCCCAACTGTCACCGGA 480

481 GCCAATGTTCAAGTGTGGGAGAAATTCGCTAGGTAATTTGAGGTTGAACCTTAAGGAA 540
 481 GCCAATGTTCAAGTGTGGGAGAAATTCGCTAGGTAATTTGAGGTTGAACCTTAAGGAA 540

541 GTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 541 GTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

601 GAGAACACCATTTGTTGGGACATCTTCTGTTCCACTTAAATGGAGAAATTCGAGAT 660
 601 GAGAACACCATTTGTTGGGACATCTTCTGTTCCACTTAAATGGAGAAATTCGAGAT 660

661 GTTAAACTCTTGAACGATCTCTTGTGCGAAAGAAACAAGAAACCGGATGGATACCA 720
 661 GTTAAACTCTTGAACGATCTCTTGTGCGAAAGAAACAAGAAACCGGATGGATACCA 720

QY 721 ATCCAGTGGATCGGCAAGTGGAGGATTCATTGCAACCCTTTTGTATCCGGAATGGAA 780
 Db 721 ATCCAGTGGATCGGCAAGTGGAGGATTCATTGCAACCCTTTTGTATCCGGAATGGAA 780
 QY 781 TGGGACTTTAGACTTCCCTTTGGTGAAGAGTATCAATGTGAGTGGTCCACAAGTATGACTT 840
 Db 781 TGGGACTTTAGACTTCCCTTTGGTGAAGAGTATCAATGTGAGTGGTCCACAAGTATGACTT 840
 QY 841 GTGTACGACGGGATTTGGTGGTGTATCTGGAGAAACAAGAGGATTTGCCTGAGGAATCTC 900
 Db 841 GTGTACGACGGGATTTGGTGGTGTATCTGGAGAAACAAGAGGATTTGCCTGAGGAATCTC 900
 QY 901 ATCTTCCATATCAATATATCTTTGGTGTACCAACCACCCTTTACTCTCAATTTCTCCAAA 960
 Db 901 ATCTTCCATATCAATATATCTTTGGTGTACCAACCACCCTTTACTCTCAATTTCTCCAAA 960
 QY 961 GFTTCAAGTCAAGTCAATGCTCAATACTACCAACTTATCCGATTTGGCCACGAGGGTTAC 1020
 Db 961 GFTTCAAGTCAAGTCAATGCTCAATACTACCAACTTATCCGATTTGGCCACGAGGGTTAC 1020
 QY 1021 AGAAATGTGATGGAGAAATTCAGAGAAATATGATCGTCTTAAGGGAAGACTTGAAGAAG 1080
 Db 1021 AGAAATGTGATGGAGAAATTCAGAGAAATATGATCGTCTTAAGGGAAGACTTGAAGAAG 1080
 QY 1081 ACAGAAAGTTCAACATCGTCTCAAGAGGAGGGAGTCCACTTGTCCGTTTCTCCITG 1140
 Db 1081 ACAGAAAGTTCAACATCGTCTCAAGAGGAGGGAGTCCACTTGTCCGTTTCTCCITG 1140
 QY 1141 AAAGATAGCAGCTGTCCACTGATTCGAAATCTCCGACATGCTTCGCAGGTATGATGG 1200
 Db 1141 AAAGATAGCAGCTGTCCACTGATTCGAAATCTCCGACATGCTTCGCAGGTATGATGG 1200
 QY 1201 ATAGTCCGCGCTTACCAATGCTCCAAATGCAACATCAATCACTGTTCTTCTGTTGGTT 1260
 Db 1201 ATAGTCCGCGCTTACCAATGCTCCAAATGCAACATCAATCACTGTTCTTCTGTTGGTT 1260
 QY 1261 ATCAGAGAGATTTCCGAGACACTCGCTGAGAGACTTGTGATCGATATAGAGAAATG 1320
 Db 1261 ATCAGAGAGATTTCCGAGACACTCGCTGAGAGACTTGTGATCGATATAGAGAAATG 1320
 QY 1321 ATCGGTGAGCTCGATGAGTTCCTTCGAGAGTATTCAAAAATATCACTTGGCAAGAG 1380
 Db 1321 ATCGGTGAGCTCGATGAGTTCCTTCGAGAGTATTCAAAAATATCACTTGGCAAGAG 1380
 QY 1381 AAGAGTGAATTAACAGCGATTAATGATGTTGATGTTGATGATGATGATGATGATGATG 1440
 Db 1381 AAGAGTGAATTAACAGCGATTAATGATGTTGATGTTGATGATGATGATGATGATGATG 1440
 QY 1441 CAGAGAGATATCATCTGGCTGGAAGAAATTTGTCGCGACGAGGAGGAGGAGGAGGAGT 1500
 Db 1441 CAGAGAGATATCATCTGGCTGGAAGAAATTTGTCGCGACGAGGAGGAGGAGGAGT 1500
 QY 1501 ATCTGCTAA 1509
 Db 1501 ATCTGCTAA 1509

RESULT 4
 BT001047 1509 bp mRNA linear PLN 18-OCT-2002
 LOCUS Arabidopsis thaliana At5g17330/MKPl1_18 mRNA, complete cds.
 DEFINITION Arabidopsis thaliana
 ACCESSION BT001047
 VERSION BT001047.1 GI:24111354
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 I (bases 1 to 1509)
 AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
 Chan,Y.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,Y.W.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.B., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis ORF clones

Unpublished
2 (bases 1 to 1509)
Chen, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.B., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (18-OCT-2002) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chan, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Chan, M.M., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, F., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.B., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source
1..1509
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U21967"
/note="This clone is in PUNI 51
ecotype: Columbia"

CDS

1..1509
/note="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
/codon_start=1
/product="At5g17330/WKPI1_18"
/protein_id="AA046801.1"
/db_xref="GI:24111355"
/translation="MWLSHAVSESDVSHVSTFASRYVTRTSLPRFMPENSIPIKAAVQ
LINDLMDGNRNLASFTVWMEPCDKLIMSSINKNYVDMDEYPTVTELRQNCVN
MIAHLFNALPEEAETAVGVTVGSSEIAMLGAFPRKQWKRKAEKGPVYDKENIYVFG
ANVQWEKFAFYPEVKLSEGYVMDPQAVDMVDENITCVAAIILGSLNGEFG
EDVLLNDLIVKMKETGWPDIHDAASGGFVAPFLYPELWDFRPLVKSINVSCH
KYGIVAGIYVWIKEDLPEELIPIHNYLGDQPTFLNFSKSSQVIAQYVQLIR
LHESYRNVENRENIVLREGLEKTERENIVSKDEGVFLVAFSLKSSCHTEFEIIS
DMLRIGNVYPAVYTPNFAHITVLRVREDFRFLAERIVDIEKVMRELDLPELFR
VIHKISLQEKSESNSDNLMTVTKKSDIKQDRDIITGMKFPVDRKTKTSGIC"

BASE COUNT

432 a 322 c 385 g 370 t
Query Match 99.9%; Score 1507.4; DB 8; Length 1509;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

1 ATGGTGCCTCCACGCCGATATCGGAGTCGGAGCTCCGTCACACTTCCACATTCGCATCA 60

Table with 4 columns: Db, Position, Sequence, and another Db. It contains genomic data for Arabidopsis thaliana, including coordinates and sequence alignments for various clones and features.

1141 AAAGATGAGCTGTCACATGAGTTCGAAATCCGACATGCTCCGACATGCTCCGAGTATGATGG 1200
 1141 AAAGATGAGCTGTCACATGAGTTCGAAATCCGACATGCTCCGACATGCTCCGAGTATGATGG 1200
 1201 ATAGTGGGGCTACACAAAGTCCAAATGCGGACAAACACATGCTTCCTGCTGGT 1260
 1201 ATAGTGGGGCTACACAAAGTCCAAATGCGGACAAACACATGCTTCCTGCTGGT 1260

1261 ATCAGAGAAGATTTCTCGAAGCACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 1261 ATCAGAGAAGATTTCTCGAAGCACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 1321 ATGCGTGAAGTTCGATGAGCTTCCTCGAAGTATTCACAAAATATCACTTGGCAAGAG 1380
 1321 ATGCGTGAAGTTCGATGAGCTTCCTCGAAGTATTCACAAAATATCACTTGGCAAGAG 1380
 1391 AAGAGTGAATTAACAGAGTAATCTGATGTCACGCTGAGAGAGCGGATATCGACAAG 1440
 1391 AAGAGTGAATTAACAGAGTAATCTGATGTCACGCTGAGAGAGCGGATATCGACAAG 1440
 1441 CAGAGAGATATCATCACTGGCTGGAAGAAAGTTTCTCGCCGACAGAAAGAGACGAGTGGT 1500
 1441 CAGAGAGATATCATCACTGGCTGGAAGAAAGTTTCTCGCCGACAGAAAGAGACGAGTGGT 1500

1501 ATCTGCTAA 1509
 1501 ATCTGCTAA 1509

AY094464 1813 bp mRNA linear PLN 05-MAY-2002
 Arabidopsis thaliana AT5g17330/MKP11_18 mRNA, complete cds.

AY094464 GI:20453186
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1813)
 Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,
 Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,
 Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
 Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,
 Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones
 Unpublished
 2 (bases 1 to 1813)
 Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,
 Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,
 Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
 Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,
 Theologis, A. and Ecker, J.R.

Direct Submission
 Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN
 Arabidopsis Full-Length cDNA) : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,
 Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E.,
 Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B.,
 Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
 Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A.,
 and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.

FEATURES
 Location/Qualifiers
 1..1813
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAF109-48-C10(R21,967)"
 /note="ecotype: Columbia"
 74..1582
 /note="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
 /codon_start=1
 /product="AT5g17330/MKP11_18"
 /protein_id="AM19834.1"
 /db_xref="GI:20453187"

CDS
 /translation="MVLSHAVSESDVSVHSTFASRVFTSLPRFKMPENSIKPEAAVQ
 IINDELMLDGNPRLNLASFTVTFMPECDKLIIMSSINKNYVDMDEYVPTTELQNRVNI
 MIAHLNPLAEATAVGVTGVSSEAIIMQALAFKRWKQKAEKGVKDPKNTVIG
 ANVQCVKPFARYFEVLEKVLSEYVMDQAVDMDVENTICVAALFSLGTLANGF
 EDVKLLDLLVEKNETGHTDIHYDAASGGIAPLPELEBWDRELPLVKINYSGH
 KGLVYAGVWIKWNRNKDIPPELPHNLGADQPTFLNFSKSSQVAFQYQILR
 LGHEHYVMNCRNINVLREGLEKTERFNIVSKDEGVPLVAFSLKSSSCHEFEFIS
 DMLRRYGVTPAYTWPNAHQHTVLRVIREDFSRLEAERLVIDIEKVMRDELDFPSR
 VIHKISLQGEKSESNLMTVTKKSDIDKQDRIITGMKFKVADRKRKTSIGC"
 BASE COUNT 504 a 384 C 417 G 508 T
 ORIGIN

Query Match 99.9%; Score 1507.4; DB 8; Length 1813;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGTCTCCACCGCGTATCGGATCGGAGTCTCGCTCCGCTCCACATCCGATCGCATCA 60
 DB 74 ATGGTCTCTCCACCGCGTATCGGAGTCTCGGAGTCTCGCTCCGCTCCACATCCGATCGCATCA 133
 QY 61 CGTTAGCTCCGACTTCACTTCTTAGTTCAGATCGCGGAAAATCGATTCCTAAGGAA 120
 DB 134 CGTTAGCTCCGACTTCACTTCTTAGTTCAGATCGCGGAAAATCGATTCCTAAGGAA 193
 QY 121 CGCGGTATCAGATCATCAACGACGAGTCTGACGGGAAATCCACGGTTGAACCTTA 180
 DB 194 CGCGGTATCAGATCATCAACGACGAGTCTGACGGGAAATCCACGGTTGAACCTTA 253
 QY 181 GCCTCTTTGTGACGACATGATGAGGACCTGAGTGTGATAAATCATATGCTCCCTCCATC 240
 DB 254 GCCTCTTTGTGACGACATGATGAGGACCTGAGTGTGATAAATCATATGCTCCCTCCATC 313
 QY 241 AACAGAATCTATGTCATGACATGACGAGTACCCCGTCAACCCGAACTTCAGAACCGATGT 300
 DB 314 AACAGAATCTATGTCATGACATGACGAGTACCCCGTCAACCCGAACTTCAGAACCGATGT 373
 QY 301 GTGAACATGATGCAATCTTAAATGACCGTGTAGAAGCGGAGACCGCGCTCGGA 360
 DB 374 GTGAACATGATGCAATCTTAAATGACCGTGTAGAAGCGGAGACCGCGCTCGGA 433
 QY 361 GTAGGACCGTGTGATCATCGGAGCCATAAATGCTGGCCGTTTGGCCCTCAAGCGTAA 420
 DB 434 GTAGGACCGTGTGATCATCGGAGCCATAAATGCTGGCCGTTTGGCCCTCAAGCGTAA 493
 QY 421 TGGCAGAAACGAGCGCAAAAGCTGAAGGCAACCCCGTCTGATAAACCACATTTGTCACCGGA 480
 DB 494 TGGCAGAAACGAGCGCAAAAGCTGAAGGCAACCCCGTCTGATAAACCACATTTGTCACCGGA 553

481 GCCAATGTTCAAGTGTCTGGGAGAAATTCGGTAGGTAATTTGAGGTTGAACTTAAAGGA 540
 554 GCCAATGTTCAAGTGTCTGGGAGAAATTCGGTAGGTAATTTGAGGTTGAACTTAAAGGA 613
 541 GTGAAATGAGTGAAGGATACTATGTGATGAGCCCTCAACAAGCTTTGATATGTTGAT 600
 614 GTGAAATGAGTGAAGGATACTATGTGATGAGCCCTCAACAAGCTTTGATATGTTGAT 673
 601 GAGAACCACTTTGTTGGGACATCTTGGTCCACCTTAACTTAACTGAGAAATCGAAGT 660
 674 GAGAACCACTTTGTTGGGACATCTTGGTCCACCTTAACTTAACTGAGAAATCGAAGT 733
 661 GTTAAACTCTTGAACGATCTTGGTCCAAAGAAACAAGAACCCGGATGGATACACA 720
 734 GTTAAACTCTTGAACGATCTTGGTCCAAAGAAACAAGAACCCGGATGGATACACA 793
 721 ATCCACGTGATCGGCAAGTGGAGATTCATTGACACCGTTTTTGTATCCGGAATTTGHA 780
 794 ATCCACGTGATCGGCAAGTGGAGATTCATTGACACCGTTTTTGTATCCGGAATTTGHA 853
 781 TGGGACTTTAGACTTCCCTTGGTGAAGATCAATGTGAGTGGTCCAAAGTATGGACTT 840
 854 TGGGACTTTAGACTTCCCTTGGTGAAGATCAATGTGAGTGGTCCAAAGTATGGACTT 913
 841 GTGATCGGAGGATTTGGTGGTGTCTGAGAAACAAGAGATTTGCTGAGGAATTC 900
 914 GTGATCGGAGGATTTGGTGGTGTCTGAGAAACAAGAGATTTGCTGAGGAATTC 973
 901 ATCTTCCATATCAATATCTTGGTGGTCCAAAGAAACAAGAGATTTGCTGAGGAATTC 960
 974 ATCTTCCATATCAATATCTTGGTGGTCCAAAGAAACAAGAGATTTGCTGAGGAATTC 1033
 961 GGTTCAAAGTCAAGTCAATGCTCAATACCAACTTATCCGATTTGGCCACGAGGTTAC 1020
 1034 GGTTCAAAGTCAAGTCAATGCTCAATACCAACTTATCCGATTTGGCCACGAGGTTAC 1093
 1021 AGAATGTGATGGAGATTCAGAGAAATATGATCGTCCATAGGAGGACTTGAAG 1080
 1094 AGAATGTGATGGAGATTCAGAGAAATATGATCGTCCATAGGAGGACTTGAAG 1153
 1081 ACAGAAAGTTCAACTGCTCAAGGAGGAGGAGTGCCTTGTGCTTCTCCTTG 1140
 1154 ACAGAAAGTTCAACTGCTCAAGGAGGAGGAGTGCCTTGTGCTTCTCCTTG 1213
 1141 AAGATGAGCTGTCACTGAGTTCGAAATCTCCGACATGCTTCGAGGATGAGATGG 1200
 1214 AAGATGAGCTGTCACTGAGTTCGAAATCTCCGACATGCTTCGAGGATGAGATGG 1273
 1201 ATAGTCCGGCTACAAATGCTCCAAATGCAACAACATCACTGCTTCGTTGGTGGT 1260
 1274 ATAGTCCGGCTACAAATGCTCCAAATGCAACAACATCACTGCTTCGTTGGTGGT 1333
 1261 ATCAGAGAAATTTCTCGAACAATCGCTGAGAGACTTGTGATGATATAGAGAAATG 1320
 1334 ATCAGAGAAATTTCTCGAACAATCGCTGAGAGACTTGTGATGATATAGAGAAATG 1393
 1321 ATGCGTGGCTGATGAGTTCCTTCCGAGATGATTCACAAAATATCACTTGGACAGAG 1380
 1394 ATGCGTGGCTGATGAGTTCCTTCCGAGATGATTCACAAAATATCACTTGGACAGAG 1453
 1381 AAGAGTGAATCTAACAGGATTAATTTGATGATGATGATGATGATGATGATGATGAT 1440
 1454 AAGAGTGAATCTAACAGGATTAATTTGATGATGATGATGATGATGATGATGATGAT 1513
 1441 CAGAGATATCATCACTGGTGGAAAGATTTGTGCTCCCAAGAGAAAGAGAGTGT 1500
 1514 CAGAGATATCATCACTGGTGGAAAGATTTGTGCTCCCAAGAGAAAGAGAGTGT 1573
 1501 ATCTGCTAA 1509
 1574 ATCTGCTAA 1582

RESULT_6
 AF353615 1776 bp mRNA linear PLN 02-MAY-2003
 LOCUS Nicotiana tabacum glutamate decarboxylase isozyme 3 mRNA, complete
 DEFINITION
 cds.
 ACCESSION AF353615
 VERSION AF353615.1 GI:13752461
 KEYWORDS
 SOURCE Nicotiana tabacum (common tobacco)
 ORGANISM
 Nicotiana tabacum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 1776)
 AUTHORS Yevtushenko,D., McLean,M.D., Peiris,S.E., Van Cauwenbergh,O.R. and
 Shelp,B.J.
 TITLE Two isoforms of tobacco glutamate decarboxylase are regulated by
 calcium/calmodulin and differ in organ distribution
 UNPUBLISHED
 JOURNAL 2 (bases 1 to 1776)
 AUTHORS McLean,M.D., Yevtushenko,D. and Shelp,B.J.
 DIRECT SUBMISSION
 TITLE Submitted (26-FEB-2001) Department of Plant Agriculture, University
 JOURNAL of Guelph, Division of Biotechnology, Bovey Building, Guelph,
 Ontario N1G 2W1, Canada
 FEATURES
 Location/Qualifiers
 1..1776
 /organism="Nicotiana tabacum"
 /mol_type="mRNA"
 /db_xref="taxon:4097"
 89..1564
 /EC_number="4.1.1.15"
 /function="converts glutamate to gamma-aminobutyric acid"
 /note="GAD; GAD4; NtGAD4"
 /codon_start=1
 /product="glutamate decarboxylase isozyme 3"
 /protein_id="AAK38667.1"
 /db_xref="GI:13752462"
 /translation="MVLKSTSESDVSHSTFASRYVTSLPRFEMAENSIPKEAIFQ
 IINDELMDGPNLNLAVFVWMEPECDKLMWDSINKNYVDMDEVPVTELOQRV
 MIARLNLAPEKETAVGVGTGSSBAIYVLAIAFKRWKQKARQKPKYKPNVTVG
 ANVQWKEFANYFEVLEKLVKGVYVMDVQAVEMVDENIYVAAILLGSFLNGEF
 EDVNLNLLLEKMKQKQWNTPIHVDVDAAGSFTAPFLYELSEWDRPLVVKKINSVGH
 KGLVYAGIENVLWTRKQDLPLEPHINFLGADQPTFLNFSKSSVVIQAYQILR
 LGVEYKNVMEKRENAI VLREGLKTRFNIVSKDEGVLVAFSLKNSRNEFEVS
 ETLRRFGWIVPAYTMDADAHVTVLWVIRDFSRFLAERLVLDIVKVLHLDLTPAR
 LSAKLEVKLVKNGKKFELEVQREVTNWKVFLARKAPVC"
 BASE COUNT 520 a 343 c 418 g 491 t 4 others
 ORIGIN
 Query Match 60.1%; Score 907.4; DB 8; Length 1776;
 Best Local Similarity 79.0%; Pred. No. 7.5e-254;
 Matches 1079; Conservative 0; Mismatches 286; Indels 0; Gaps 0;
 QY 1 ATGTTGCTCTCCACGCGGATCGGAGTGGAGCTTCGCTCCACATCCACATTCGCATCA 60
 DB 89 ATGTTTCTTCGAAACCTCTTGAGTGGACGCTTCGGTACACCTTCCTTTCCTCT 148
 QY 61 CGTTCACGTCGCTTCTTCACTTCCTTAGGTTCAAGATCCCGGAAACTCGATTCCTAAGGAA 120
 DB 149 CGCTATGTTCCAACTTCCCTTCCAAAGTTCAGATGGCGGAGAAATTCGATACCAAAGAG 208
 QY 121 GCGCGTATCAGATCATCACGACGAGCTGATCTTACGCGGAATCCACGGTTGACITTA 180
 DB 209 GCGGCATTTCAAATAATTAACGACGATTTGANGCTTTCGCGGAAATCCAAAGGCTGAACTT 268
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATGTCTCCCATC 240
 DB 269 GCTTCATTTGTGACCAATGATGGAGCCAGAGTGTGATAAGCTTATGATGGACCTCAT 328
 QY 241 AACAGAACTATGTTGACATGGACGAGTACCCCGTCCACCAACCGAACTTCAGAACCGGAT 300

241 AACAAAGAACTATGTTGACATGGAACGAGTACCCCGTCCACCACCGAACTTCAGAACCGATGT 300
 311 AACAAAGAACTATGTTGACATGGAACGAGTACCCCGTCCACCACCGAACTTCAGAACCGATGT 370
 301 GTGAACAATGATTCGACATCAATTCATGACACCGTTAGAAAGCGGAGACCCCGTCCGGA 360
 371 GTAAACATGATAGCTCATTTGTTTAAACCGACCCATTTGGAGATGGAGAGACTCGATTTGA 430
 361 GTAGGAACCGTTGGATCATCCGAGGCGCAATAAGTTGCGCGGTTTGGCCCTCAAGGTA 420
 431 GTTGGAACTGTTGGATCCCTGAGGCTATATGCTTGGATTAGCTTCAAGAGAAA 490
 421 TGGCAGAACAAAGCCGAAGCTGAAGGCAAAACCGGTGCGATAAACCACCAATGTCACCGGA 480
 491 TGGCAAAATAAATAAGAACCCCAAGCAAGCCCGTGTGACAAAGCCCAATAATGTCACCTGT 550
 481 GCCAATGTTCAAGTGTGTTGGGAAATTCGCTAGGTAATTTGAGGTTGAACTTAAAGNA 540
 551 GCCAATGTCAGGTTGTTGGGAAATTTGGAGGTAATTTGAGGTTGAACTTAAAGNA 610
 541 GTGAAATGATGAGGATCACTATGATGAGACCTCAACAAAGCTGTTGATGTTGAT 600
 611 GTAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
 601 GAGAACCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGAT 660
 671 GAGAACCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGAT 730
 661 GTTAAACTCTTGAACGATCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
 731 GTTAAAGCTTGAATGACCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 790
 721 ATCCAGTGGATCGGCAAGTGGAGATCAATGACCCGTTTATCCGGAATGGAA 780
 791 ATTCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850
 781 TGGACCTTTAGACTTCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
 851 TGGACCTTTAGACTTCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 910
 841 GTGTACGAGGATGGTGGTGAATGTTGAGAAACAAAGAGGATTTGCTGAGGAACTC 900
 911 GTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970
 901 ATCTTCAATATCAATATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
 971 ATCTTCAATATCAATATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1030
 961 GGTTCAGTCAAGTCACTGCTCAATACCTCAACTTATCCGATGTTGGCCACGAGGTTAC 1020
 1031 GGTTCAGTCAAGTCACTGCTCAATACCTCAACTTATCCGATGTTGGCCACGAGGTTAC 1090
 1021 AGAAATGATGAGAAATTCGAGAGAAATGATCGTCTTAAGGAAAGACTTGAAG 1080
 1091 AAGAAATGATGAGAAATTCGAGAGAAATGATCGTCTTAAGGAAAGACTTGAAG 1150
 1081 ACAGAAAGTTCAACTGCTCAAGGAGGAGGATGCTCACTTCTGCTTCCCTTGG 1140
 1151 AGTGAAGATTCACATCAATCAAGAAATTTGGAGTTCCATGATGATGATGATGATGATGAT 1210
 1141 AAAGATGAGCTGTCTCACTGATTTGAAATTCGACATGCTCCGACATGCTCCGAGTATGATGG 1200
 1211 AAAGATGAGCTGTCTCACTGATTTGAAATTCGACATGCTCCGACATGCTCCGAGTATGATGG 1270
 1201 ATAGTGGCGCTACAAATGCTTCAAAATGCAACACATCACTGTTCTTCTGTTGTTGTT 1260
 1271 ATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330
 1261 ATCAGAAAGATTTCTCGAGAACTCGCTGAGAGACTTGTGATGATGATGATGATGATGATGAT 1320
 1331 ATTAGAGAAATTTCTCCGCTACACCTCCGAGGACTGGTAAATGACATGTAAGAAAGTIC 1390
 1321 ATGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380

1391 CTCACGAGCTAGACACACTTCCGCGAGGCTCAACGCTAAGCTAGCCGTCGCGAGGCG 1450
 1381 AAGAGTCAATCTAACAGCGATACTTGTGATGCTCACCGTGAAGAGAGCGATATCGACAG 1440
 1451 AATGG-----CAGCGGCTGCATAAAGAAAACAGATAGAGAAGTG 1489
 1441 CAGAGATATCATCACTGCTGGAAGAAATTTGTCGCCGACAGGAAAGACGACT 1497
 1490 CAGCTTGAGATTACTACTGCAATGGAAGAAATTTGTCGATGATAAAGAAAGAAAGACT 1546

RESULT 8
 PETGAD
 LOCUS Petunia hybrida glutamate decarboxylase (gad) mRNA linear PLN 15-OCT-1993
 DEFINITION Petunia hybrida glutamate decarboxylase (gad) mRNA, complete cds.
 ACCESSION L16797
 VERSION L16797.1 GI:294111
 KEYWORDS glutamate decarboxylase.
 SOURCE Petunia x hybrida
 ORGANISM Petunia x hybrida
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; lamiales; Solanales; Solanaceae; Petunia.
 REFERENCE 1 (bases 1 to 1785)
 AUTHORS Baum, G., Chen, Y., Arazi, T., Takatsuji, H. and Fromm, H.
 TITLE A plant glutamate decarboxylase containing a calmodulin binding domain. Cloning, sequence, and functional analysis
 J. Biol. Chem. 268 (26), 19610-19617 (1993)
 JOURNAL J. Biol. Chem. 268 (26), 19610-19617 (1993)
 MEDLINE 93374956
 PUBMED 8366104
 COMMENT Original source text: Petunia hybrida flower corolla cDNA to mRNA.
 FEATURES
 Location/Qualifiers
 1..1785
 /organism="Petunia x hybrida"
 /mol_type="mRNA"
 /db_xref="taxon:4102"
 /tissue_type="corolla"
 /dev_stage="flower"
 gene
 1..1785
 /gene="gad"
 CDS
 72..1574
 /EC_number="4.1.1.15"
 /codon_start=1
 /product="glutamate decarboxylase"
 /protein_id="AAA33709.1"
 /db_xref="GI:294112"

BASE COUNT 550 a 298 c 401 g 536 t
 ORIGIN
 Query Match 58.4%; Score 881.8; DB 8; Length 1785;
 Best Local Similarity 74.9%; Pred. No. 2.3e-246;
 Matches 1121; Conservative 0; Mismatches 367; Indels 9; Gaps 1;
 Qy 1 ATGCTGCTCCACGCGGATCGGAGTGGAGCTCCGTCACCTCCACATTCGATCA 60
 Db 72 ATGCTTCTAACAAGACAGTGTCCGACGCGATGTGTCCATTCACTCCAGTTTGTCTT 131
 Qy 61 GGTTCAGTCCGCTACTTCCATTCCTAGGTTCAAGATGCGGAAACTCGATTCCTAAGGAA 120
 Db 132 CGATATGTTGAACTTCTTCCAGGTTTAAATGCCAGATATTCGATACCAGAAAGAA 191
 Qy 121 GCGCGTATCAGATCATCAACGACGAGCTGATGCTTTCAGCGGAAATCCAGGTTGAATTA 180

/protein_id="AAK18620.1"
/db_xref="GI:13310813"
/translatlon="MVLSTKASBDSVSIHSTFASRYVYRSLPRFKMPENSIPKEAAAYQ
IINDELMDGNPRMLNLFASFVWMEPCNKLMDSINKNYVDMGEYPTVTELRQNRV
MIAHLFNAPLGDGTAVGTVGSSBAIMLAGLAFKWKONKMAQCKDKPNIYTG
ANVQVCKEAPRYEYVAVKLSGSDYVWDPKAVEMWENTICVAALLGSLNIEGF
EDVRRDLLEKRNKFTGMDTPIHVDRAAGSFAPFLYPELWDFRPLVKSINVSGB
KYGLVYAGIOWAIMWRNKEDLPDELIFHNYLGDQPTFLNFKSGSSQVIAQYQYLI
LGFEGYKNVMEQENARVLEKLEKSRFNIIISKEIGVPLVAFSLKMSQHNFEFIS
ETLRFQWIIIPAYTWPNAQHVIVRVVIREDFSRTEAERLVIDIIEKVLHELDLTPAR
VNAKLVAEANGSVHKKTDREVOLEITAAWKFVADKDKKTINGVC"
SE COUNT 511 a 280 c 392 g 483 t 6 others
IGIN

Query Match 58.1%; Score 877; DB 8; Length 1672;
Best Local Similarity 74.9%; Pred. No. 5.9e-245;
Matches 1121; Conservative 0; Mismatches 355; Indels 21; Gaps 1;

- 1 ATGGTCTCCCGCGTATCGAGTGGAGCTCCGTCCTCCACATCGCATCA 60
- |||||
- 4 ATGGTCTCGTCCAGACAGCGTGGAAAGTACCGTCCATCCACTCCACTTC 63
- |||||
- 61 CGTTACGTCCTACTTCACTTCCCTAGGTTCAAGATGCCGGAATACTCGAATCCTAAGGAA 120
- |||||
- 64 CGATATGTTGGAATCTTCCAGGTTTAAAGTCCAGAGAATTTCTATACCMAAGAA 123
- |||||
- 121 CGGGGTATCATGATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
- |||||
- 124 GCAGCAATCAAAATCAAAATCATGAGCTTATGTTAGATGGAATCCAAAGCTAAATTTA 183
- |||||
- 181 GCCTCCTTTGGACGACATGATCGAGCTGAGTGTGATTAACACTCATCACTCCTCCATC 240
- |||||
- 184 GCATCTTTTGTGACAAATGATGGAACAGAGTGTAAATAGTTAAATGATGGATTCAT 243
- |||||
- 241 AACAAAGACTATGTTGATGATGACGAGTACCCCGTCCACCGAATTCAGAACCGATGT 300
- |||||
- 244 AACAAAGACTAGTTGATGATGAGTGAATFACCTGTAAACACTGAGCTTCAGAAATCGATG 303
- |||||
- 301 GTGAACATGATGACATCTATTCATGACCCGTTAGAAAGGCGGAGACCGCGTGGGA 360
- |||||
- 304 GTAAATATGATAGTCAATTTGTTAAACGACCACTTGGAGATGGAGAGACTGCAAGTTGGA 363
- |||||
- 361 GTAGAACCGTTGATGATCATCGAGGCGCATATGTTGGCCGTTTGGCCTCAAGCGTAAA 420
- |||||
- 364 GTTGAACGTGTTGGATCTCGAGGCTTATGCTTGTGTTGATGATGATGATGATGATGAT 423
- |||||
- 421 TGGCAGAACAGCGCCAAAGCTGAAGCAGAACCCCGTTCGATTAACCCCAACTTGTCCCGGA 480
- |||||
- 424 TGGCAAAATAAAATGAAAGCCMAGGGAAGCCCTGCGACAAAGCCCAATATGTCACCTGGT 483
- |||||
- 481 GCCAATGTTCAAGTGTGGAGAAATTCCTAGTACTTTGAGTTGAACTTAAGGAA 540
- |||||
- 484 GCCAATGTTCCAGGTTGTTGGAGAAATTTCCAAAGTATTTTGAAGTGGAGTTGAAAGAA 543
- |||||
- 541 GTGAAATGAGTGAAGGACTATGATGATGACCTCAACAAAGCTTGTGATGATGATGATGAT 600
- |||||
- 544 GTAAAATGAGTGAAGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 603
- |||||
- 601 GAGAACACCAATTTGTTGCGGACATTTTGGTCCACTTTAATGGAGAATTCGAAGAT 660
- |||||
- 604 GAGAAATCAATTTGTTGATGCTACTTCTGGTTCCTCACTTCAATGGTGAATTTGAAGAT 663
- |||||
- 661 GTAAACTTTGAAAGACTTCTTGTGAAAGAACAAAGAACCGGATGGGATACCA 720
- |||||
- 664 GTTAAAGCGCTTGAATGACTTCTTGTATGAGAAAGAACAAAGAACCGGATGGGACACTCCA 723
- |||||
- 721 ATCCACGTTGATGCGGCAAGTGGAGATTCATTTGACCGTTTTTGTATCCGCAATTTGAA 780
- |||||
- 724 ATTATGTTGATGACAGAGTGGTGGATTCATTTGACCACTTCTTTATCTTACCTTGGAT 783
- |||||
- 781 TGGACTTTAGACTTCCCTTGGTGAAGATATCAATGAGTGTGTCACAGATGGACTT 840
- |||||
- 784 TGGGATTTTAGATTTACCAATTTGGTGAAGATTTAATGATGATGATGATGATGATGATGATGAT 843
- |||||

- 841 GTTAGCGAGGAAATGGTGGTGTCTGGAGAAACAAGAGGATTTCCCTGAGGAATC 900
- |||||
- 844 GTCTATGCTGTTATGGTGGCCATTTGGAGAAATAAGGAAGACTTGGCTGATGAACTT 903
- |||||
- 901 ATCTTCCATATCAATTAICTTGGTCTGACCAACCACCTTTACTCTCAATTTCTCCAAA 960
- |||||
- 904 ATTTTCCACATTAATTAICTTGGTCTGATCAACCTACTTTTCCACTCTCAACTTCTTAAA 963
- |||||
- 961 GGTTCAAAGTCAAGTCAATCTCAATTAICTTCCAACTTATCCGATTTGGCCACGAGGTTAC 1020
- |||||
- 964 GGTTCAGCAAGTAAATGCTCAATTAICTTCCAACTTATTCGCTTGGGTTTGGAGGTTAC 1023
- |||||
- 1021 AGAATGTTGATGAGAAATGTCAGAGAAATATGATGCTTAAGGGAAGCACTTGAGAA 1080
- |||||
- 1024 AAGAATGTTATGAGAAATGTCAGAAATAAGAGGTAAGAGAAAGGACTTGAATAA 1083
- |||||
- 1081 ACAGAAAGGTTCAACATCGTCTCAAGGAGCAGAGGAGTCCCACTTGTGCTTTCCTTTG 1140
- |||||
- 1084 AGTGGAGAAATCAATTAICTTCCAAAGAAATGGAGTTCCATTTAGTAGCTTTTCTCTCT 1143
- |||||
- 1141 AAGATGATGAGCTGTCACTGAGTTGAAATCTCCGACATGCTTCGAGATATGGATGG 1200
- |||||
- 1144 AAGACAACAGTCAACACAAATGAGTTGCAAAATTTCTGAAACTTTAGAAATTTGATGG 1203
- |||||
- 1201 ATAGTCCCGCTTACAAATGCTTCAAAATGCAACAACAATCACTGTTTCTGTTGTTGTT 1260
- |||||
- 1204 ATTAATTTCTGCAATATGATGCTTCAACCAAAATGCTCAACATGTCACAGTTTCTCAG 1263
- |||||
- 1261 ATCAGAAAGATTTCTCGAGAACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGT 1320
- |||||
- 1264 ATTAGAGAAATTTCTCCGCTACACTCCGCGGAGCAGCTGTTGATGACATTTGAAAAGT 1323
- |||||
- 1321 ATGCTGAGCTCGATGAGCTTCTTCGAGAGTATTCACAAAATATCACTTGGCAAGAG 1380
- |||||
- 1324 CTCACAGCTAGACACACTTCCGGGAGGTTCAACGGTAAAGCTCCGCTGGCCGAGGCG 1383
- |||||
- 1381 AAGATGAAATCAACAGGATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1440
- |||||
- 1384 AATGG-----CAGCGCGTGCATTAAGAAAACAGATAGAGAAAGT 1422
- |||||
- 1441 CAGAGAGATATCATCACTGGCTGGAGAAATTTGTCGCGCAGGAAAGAGAGAGT 1497
- |||||
- 1423 CAGCTGAGATTAATGCTGATGAGAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1479
- |||||

RESULT 12
AY124873
LOCUS 1485 bp mRNA linear PLN 06-JUL-2002
DEFINITION Arabidopsis thaliana At1g65960/F12P19_12 mRNA, complete cds.
ACCESSION AY124873
VERSION AY124873.1 GI:21700916
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
s.
REFERENCE 1 (bases 1 to 1485)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shimm,P., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1485)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shimm,P., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,

Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.
 Direct Submission
 Submitted (17-JUN-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, F., Karlin-Neumanna, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers
 1. 1485
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="UI4793"
 /note="This clone is in pUNI 51 ecotype: Columbia"
 1. 1485

/note="glutamate decarboxylase 2"
 /codon_start=1
 /product="Atlg65960/F12P19_12"
 /protein_id="AA070582.1"
 /db_xref="GI:21700917"
 /translation="MVLTKATIDNDESVCTMFGSRVRLPLKYEIGENISPKDAAYQI IKLEMDGPRMLNLSAFVTTWMEPECKLIMDSINKYVYDMEYPTVTELOQRVNI IARLFLAPLEESFETAVGVTVGSEYVMDPKAEMVDENTICVAALIGSTLNGFE NVQVWEKFAFYFEVLEKNSGYYVMDPKAEMVDENTICVAALIGSTLNGFE DVKRLNDLLVKNBEETGWNTPHVDVAASGGFIAPFIYPELWDFRPLPLVKSINVGHF YGLVYAGIGWVMAEEDLPEELIHFHNYLGAQDPTFLNFKSSQIQAQYVQLIHL GFEGYKVMNENCIENMNVYVLEKIEKTEFNIVSKDQGFVVAFLSKDHSFHNFEISE MLRFGMIVFAYTMDPAQHTVLRVREDRDFSKLAEKRLVADISKVLHDLTFLPSKI SKRWGIEGIAENVKEKKEKLEMEVIVGWRKFKKRWKQVGC"
 SE COUNT 438 a 290 c 385 g 372 t
 IGIN

Query Match 57.7%; Score 870.2; DB 8; Length 1485;
 Best Local Similarity 75.1%; Pred. No. 5.6e-243;
 Matches 1110; Conservative 0; Mismatches 348; Indels 21; Gaps 1;
 31 GACGTCCTCCGTCACCTCCACATTCGGCATCAGTTCAGTTCGGTCTTCCATTCCTAGGTTTC 90
 28 GATGAATGTCGACCATGTTGGATCTCGCTATGTTCCACTACACTCCCAAGTAT 87
 91 AAGATGCGGAAACTGATTCCTAAGAAAGCGCGGTATCAGATCATCAACGACGAGCTG 150
 88 GAGATGTTGAGAAATTCGATACCGAAAGCGCTCGGTATCAGATCATATAAAGATGAGCTG 147
 151 ATGCTTGACGGGATCCACGGTTCGAACTTAGCTCCCTTTGTCGACGATGATGAGCTT 210
 148 ATGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
 211 GAGTGTGATAAACAATCATGCTCCATCAACAAGAACTATGTTGACATGAGGAGCTAC 270

Db	208	GAGTGTGACAAACTCATGAGTCTTATCAATAGAACTACTAGTTCGATATGATGATGAGTAC	267
QY	271	CCCGTCCACCACCACTTCAGAACCCGATGTGAACATGATTCACATCTATTCAATGCA	330
Db	268	CCCTGTCCAACTGAGCTCCAGAACCGATGTGAAACATTTATAGCTCGACTGTTCAATGG	327
QY	331	CCGTTAAGAGAGCGGAGACCCCGTTCGAGTAGGAACCGTTGGATCATCGGAGGCCATA	390
Db	328	CCACTCGAGAAATCTGAGACGGCGGTGGAGTAGGACAGTTCGTTTTCAGAAAGCAATC	387
QY	391	ATGTTGGCGGTTTGGCCCTTCAAGCTTAATGCGAACAAGCGCAAGCTCAAGGCAAA	450
Db	388	ATGTTAGCCGGATTTGGCCCTTCAAAGAAAATGCGAACAAGCGCAAGCTCAAGGTA	447
QY	451	CCCGTTCGATATAAACCACCAATGTCACCGGAGCCAAATGTTCAAAGTGTGTTGGGAAATTC	510
Db	448	CCCTATGACAAACCAACATTTGTCACCGGAGCAATGTTCAAAGTTCGTTGGGAAATTC	507
QY	511	GCTAGTACTTTGAGGTTGAACTTAAGAAAGTGAATTCAGTGAAGGATACTATGTGATG	570
Db	508	GCTCGGTACTTCGAGTGGAGTAAAGGAAAGTAAACCTTAAGTGAAGGTTACTACGTTG	567
QY	571	GACCTCAACAAGCTGTTGATATGTTGATGAGAACACCAATTTGTTGCGGACATTC	630
Db	568	GATCCAGACAAAGCAGCAAAATGTTAGACGAGAACACAACTGTTGTCGACCCATATTC	627
QY	631	GGTTCACCTTTAATGGAGAATTCGAAAGATGTTAACTCTTGAACGATCTCTTGGTTCGAA	690
Db	628	GGATCCACACTCAACGGTGAATTCGAAAGCGTGAACCGTCTCAATGACTTGTAGTCAAG	687
QY	691	AGAACAAAAGAACCGGATGCGGATACCAACCTCCAGTGGATCCGCAAGTGGAGATTC	750
Db	688	AAAAACGAGGAGACTGGTTGGAAACACACCGATCCACGTTGGATGCAAGTGGAGGTTTC	747
QY	751	ATTCACCCCTTTTGTATCCGGAATTTGGAATGGGACTTTAGACTTCCCTTGGTGAAGCT	810
Db	748	ATAGCTCCGTTTATCTATCTGAAATTTAGAAATGGGACTTTAGGCTTCCCTTGGTTRAGAC	807
QY	811	ATCAATGTGAGTGGTCCAAAGTATGGACTTGTGTAGCGAGGATGGTTGGTGTGATCTGG	870
Db	808	ATCAAGCTGAGTGGTCCAAAGTATGGACTTGTGTAGCGAGGATGGTTGGTGTGATCTGG	867
QY	871	AGAACAAAAGAGATTTGCTGAGAACTCACTTCCATATCAATTAATCTTGGTCTGATC	930
Db	868	AGGGCAGCAGAGATTTACTGAAAGCTTATCTTTCAATTAATTAATCTTGGTCTGAT	927
QY	931	CAACCCACCTTTACTCTCAATTTCTCAAAGTTCAAAGTCAAGTCAATCTCAATCAACTAC	990
Db	928	CAACCCACCTTTACTCTCAATTTCTCAAAGTTCAAAGTCAAGTCAATCTCAATCAACTAC	987
QY	991	CAACTTATCCGATTTGGCCACCGAGGTTACAGAAATGTGATGGAGAAATTCGAGAGAA	1050
Db	988	CAGCTCATTCGCTTTGGATTCGAGGGGTACAAAATGTTGATGGAGAAATTCATAGAGAA	1047
QY	1051	ATGATCGTCTTAAGGAGGACTTNGAAGAGCAGAAAGTTCAAACATCGTCTCAAGGAC	1110
Db	1048	ATGGTGGTTCCTCAAAAGAGGTATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAGGAC	1107
QY	1111	GAGGAGTGCACCTTTCGCTTTCTCTTTGAAAGATGACGAGCTGTCCACATGAGTTCGAA	1170
Db	1108	CAAGGATGCCAGTTCGCTTCTCTCTCAAGGACCATAGTTTCCACAAACGAGTTCGAG	1167
QY	1171	ATCTCCGATGCTGCGAGGTTATGATGAGTGTAGTCCCGGCTTACAAATGCTCCAAAT	1230
Db	1168	ATCTCTGAGATGCTACCTGCTTTTGGCTGGATCGTCCAGCTTACACTATGCTCCGAT	1227
QY	1231	GCACAAACATCACTGTTCTTCTGTTGTTTACAGAGAGATTTCTCGAGACATCGCT	1290
Db	1228	GCACAGCATCACGTTCTGCTGTTGTTGTCATCAGGAAAGATTTCTCAAGAAACATCGCG	1287
QY	1291	GAGAGACTTGTGATCGATATAGAGAAAGTGTGCGGTGAGCTCGATGAGCTTCTTCGAGA	1350
Db	1288	GAGAGACTTGTGCTGATATTCGAAAGGTTCTCATGAGCTAGATACCTTCCCTTCAG	1347

1351 GTGATTCACAAATATCCTCTGGACCAAGAGAGAGTGAATCTAACACCGGTAACCTTGATG 1410
 1348 AT-----ATCTAAGAAAGATGGAAATAGAAAGGATCGCGGAAAT 1386
 1411 GTCACGGTGAAGAGAGGATATCCACAAGCAGAGAGATATCATCACTGGTGGGAAGAG 1470
 1387 GTAAGGAGAGAGATGGAGAGAGATCTGTATGGAAGTATTGTTGGATGGAGGAAG 1446
 1471 TTTGTCCGACAG 1509
 1447 TTTGTGAAG 1485

SULT_13
 U46665
 CUS 1676 bp mRNA linear PLN 01-SEP-1998
 Arabidopsis thaliana glutamate decarboxylase 2 (GAD2) mRNA,
 complete cds.
 U46665
 U46665.1 GI:1184959
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 1676)
 Turano, F.J. and Fang, T.K.
 Characterization of two glutamate decarboxylase cDNA clones from
 Arabidopsis
 Plant physiol. 117 (4), 1411-1421 (1998)
 93369120
 9701597
 2 (bases 1 to 1676)
 TURANO, F. J. and THAKKAR, S.
 Direct Submission
 Submitted (21-JAN-1996) Frank J. Turano, Climate Stress Lab,
 USDA/ARS, BARC-West, Bldg. 046A, Beltsville, MD 20705, USA

location/Qualifiers
 1. .1676
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="ecotype Columbia"
 /db_xref="taxon:3702"
 1. .1676
 /gene="GAD2"
 /name="GAD2"
 /codon_start=1
 /product="glutamate decarboxylase 2"
 /protein_id="AAC33485.1"
 /db_xref="GI:1184960"
 /translation="MVLTKTANDSEVYTFMGRSRYVRLTLPKYIEIGNSIPKDAAYQI
 IKBLMLDGNPRNLASFVTVMEPCDKLIMDSINKVYVMDPEYFVYVTELOLNKRVNI
 LARLFNAPLESETAVGVGVSSSEALMLAGLAFKRWKONKRAKGPYKPKNIVTGA
 NVQVCEKARYEVELKENVLSRGYVMDPKNAEMVDENITCVAAILGSLNGEPE
 DVKRLNLLVKNEBEGNTPIHVDAAAGGFIAPIYPELEWDFRFLPLVKSINVSGHK
 YGLVYAGIWMVRAEDELPEELFFHINYLGAQPTFLNFKSGSQIIAQYYQLRL
 GFEGYKNVMEINMVLKKEIKETRFNIVSKDQGVVAFSLKDHSHFNEFISE
 MLRFQWIVFAYTMPADAQHTVLRVVRIDRFRTLAERLVADI SKVLHDLTLPSKI
 SKKMGIEIAENVEKMEKELMEIVGVWRKFKYKRRKMGVC"
 498 a 313 c 424 g 441 t

gene
 CDS
 Query Match 57.7%; Score 870.2; DB 8; Length 1676;
 Best Local Similarity 75.1%; Pred. No. 5.7e-243;
 Matches 1110; Conservative 0; Mismatches 348; Indels 21; Gaps 1;
 31 GACGTCCTCCACTCCACATTCGGATACGTTACGTCGCGTACTTCACTCCPAGGTTTC 90
 44 GATGAATCTGTCTGCACCAATGTCGGATCTCGCTATGTTCCGACTACACTCCCAAGTAT 103

QY 91 AAGATCCCGGAAAACTCGATTCCTTAAGGAAGCGCGGTATCAGATCATCAACGACGAGCTG 150
 DB 104 GAGATTTGGTGAAGATTCGATACCGAAAGACCGTGCATATCAGATCATATAAAGATGAGCTG 163
 QY 151 ATGCTTACCGGAAATCCACCGTTGAACCTTAGCCTCCTTTGTGACGACATGGATGGAGCCT 210
 DB 164 ATGCTTGAATGGTAAACCGGAGGCTTAACCTAGCTTGGTTTGTGACTACATGGATGGAACA 223
 QY 211 GAGTGTGATAAATCTCATCATGCTCCATCAACAAGAACTATGTTGACATGACGACGATAC 270
 DB 224 GAGTGTGAAAACTCATCATGACTCTATCAACAAGAACTATGTTGATGATGGATGAGTAC 283
 QY 271 CCGGTCACCAACCGAATTCAGAACCGATGTTGACATGATTTGCACATCTATTTCAATGCA 330
 DB 284 CTTGTCAACAACCTGAGCTCCAGAACCGATGTTAAACATTTATAGCTCGACTGTTCAATGG 343
 QY 331 CGTGTAGAAGAGCGGAGACCCCGTGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
 DB 344 CCACCTGAGGAATCTGAGACCGGCTGAGGAGTAGGACAGATTTGGTTCTTCAGAAAGCCATC 403
 QY 391 ATGTTGGCCGGTTTGGCTTCAAGCGTAAATGGCGAAACAAGCGGAAAGCTGAAGCAAAA 450
 DB 404 ATGTTAGCCGGATTTGCCCTTCAAAAAGAAAAATGGCAGAAACAAAACGCAAGGCTGAGGTA 463
 QY 451 CCGGTCGATAAACCACCAATTCACCGGAGCAATGTTCAAGTGTGTTGGGAGAAATTC 510
 DB 464 CCTATGACAAACCCCAACATTTGCTAGGAGCAATGTTCAAGTGTGTTGGGAGAAATTC 523
 QY 511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTCAAATTTGAGTCAAGGATCACTATGTGATG 570
 DB 524 GCTCGGTACTTTCGAGGTGAGCTAAAGGAAGTAAACCTTAAGTGAAGTTACTACCTGATG 583
 QY 571 GACCTCAACAAGCTGTTGATGATGTTGATGAGACACCAATTTGTTGTCGGACATTCCT 630
 DB 584 GATCCAGACAAAAGCAGCAAAAATGTTAGCAGAGACACAAATCTGTGTCAGCCATATG 643
 QY 631 GPTTCCACTCTTAATGGAGAAATTCGAAGATGTTAAACTCTTCAACGATCTCTTTGTCGAA 690
 DB 644 GGAATCCACTCAACCGTGAAGTTCGAAGACGTTGAAGACGTTCAATGACTTGTGTAGTCAAG 703
 QY 691 AAGAAACAAAAGAACCCGATGGGATACCAACATCCACGTCGATCGCGCAAGTGGAGATTC 750
 DB 704 AAAAAAGGAGAGACTGTTGGAACACACACCGATCCACGTCGATGACGAAATGGAGGTTTC 763
 QY 751 ATGCAACCGTTTTTGTATCCGGAATGGAATGGACTTTAGACTTCCCTTGGTGAAGAT 810
 DB 764 ATAGCTCCGTTTTATCTATCTGAAATTTAGAATGGACTTTAGACTTCCCTTTGGTTAAGAT 823
 QY 811 ATCAATGTGAGTGGTCAACAAGTATGGACTTGTGACGAGGATGGTTGGTGTGATCTGG 870
 DB 824 ATCAACGTTGAGTGGTCAACAAGTATGGACTTGGCTATGCTGTTGTTGGTTCGTTGG 883
 QY 871 AGAAACAAAAGAGAAATTTGCTGAGGAACCTCATCTTCCATATCAATTTATCTTGGTGTGAC 930
 DB 884 AGGGCAGCAGAGAAATTTGCTGAAGAGCTTACTCTTCAATTAATTAATCTTGGTGTGAT 943
 QY 931 CAACCCACTTTACTCTCAATTTCCCAAAGTTCAAGTCAAGTCAATGCTCAAACTACAC 990
 DB 944 CAACCCACTTTACTCTCAATTTCCCAAAGGATCGAGCAAAATTTTGTCTCAACTACTAC 1003
 QY 991 CAACTTATCCGATTTGGCCACGAGGTTACAGAAATGTGATGGAGAAATTCGACAGAGAA 1050
 DB 1004 CAGCTCATTTGCTTTGGATTCGAGGGGTACAAAATATGATGGAGAAATTCATAGAGAAC 1063
 QY 1051 ATGATCGTCTTAAGGAAGGACTTGAAGACAGAAAGGTTCAACATCGTCTCAAAAGGAC 1110
 DB 1064 ATGTTGTTTCTCAAAGAGGATAGAGAAAACAGAGCTTTTCAACATAGTCTCAAAAGGAC 1123
 QY 1111 GAGGAGTCCACTGTCGCTTTCTCCTTGAAGATGACAGCTGTCAACTGAGTTCGAA 1170
 DB 1124 CAAGGAGTCCACTGTCGCTTTCTCTCAAGGACCAATGTTTCCAAACAGGATTCGAG 1183
 QY 1171 ATCTCCGACATGCTTCGAGGTATGGATGATAGTTCGCGGCTCACAAATGCGCTCCAAT 1230

1184 ATCTGTGATGCTACGTCGTTGGTGGTGGATGTCACGCTTACACTATGCTCCGCGGAT 1243
 1231 GCACAACATCACTGTTCTTCGTGGTATACAGAGAATTTCTCGAACAACACTCGCT 1290
 1244 GCACAGACATCACGGTTCGCGTGTGTCATCAGGGAAGATTTCTCAAGAACAACCTCGC 1303
 1291 GAGAGACTGTGATCGATATAGAGAAGTGAATGCGTGGAGCTGATGAGCTTCTTCGAGA 1350
 1304 GAGAGACTGTTGCTGATATTTCAAGAGTGTTCATGAGCTAGATACCTTGCCTTCCAAG 1363
 1351 GTGATTCACAAAATFATCACTTGGCAAGAGAGATGAATCTAAACACGATTAACCTTGATG 1410
 1364 AT-----ATCTAAGAGATGGGATAGAGAGGATCGCGGAAAT 1402
 1411 GTCAGGTGAAGAGAGCGATATCGCAAGCAGAGAGATATCACTACTGGCTGGAAGAAG 1470
 1403 GTAAGGAGAAGAAGATGGAGAAGAGATTTCTGATGGAAGTATTTGTTGGATGGAGGAA 1462
 1471 TTTGTCGCCGACAGAGAAGACAGAGTGGTATCTGCTAA 1509
 1463 TTTGTGAAGGAGGAGAAGATGAATGGTGTGTGCTAA 1501

SULT 14
 U49937 1681 bp mRNA linear PLN 13-AUG-1998
 Arabidopsis thaliana glutamate decarboxylase (GAD2) mRNA, complete cds.

U49937 GI:1236618
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1681)
 Zik, M., Arazi, I., Snedden, W.A. and Fromm, H.
 Two isoforms of glutamate decarboxylase in Arabidopsis are regulated by calcium/calmodulin and differ in organ distribution
 Plant Mol. Biol. 37 (6), 967-975 (1998)

98363649
 9700069
 2 (bases 1 to 1681)
 Zik, M., Arazi, I. and Fromm, H.
 Direct Submission
 Submitted (25-FEB-1996) Hillel Fromm, Plant Genetics, Weizmann Institute of Science, Rehovot 76100, Israel
 Location/Qualifiers
 1. .1681
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="5B2T7B"
 /note="using EST found in GenBank Accession Number T04804#"

gene 1. .1681
 CDS 1. .1485
 /gene="GAD2"
 /function="calcium/calmodulin binding"
 /note="glutamic acid decarboxylase; GAD; similar to A. thaliana GAD encoded by GenBank Accession Numbers U10034 and U45665"
 /codon_start=1
 /product="glutamate decarboxylase"
 /protein_id="AAC31617.1"
 /db_xref="GI:1236619"
 /translation="MVLTKTNDSEVCTMFGSRVITLLPKYIEIGENSIPKDAAYOI IKDELMDGNPRNLASGVITWMEPCDKLIMDSINKNVYDMDEYPTVTELQNRKCVNI IARLFNAPLEBSEATAVGVIGVSSEAIMLAGLAPKRWKRAEKPKYKPNIVTGA

NTVQCWEKPYRYFVELKEVNLSEGYVMDPKAAEMVDENIICVAAILLSTLNGEFFE
 DVKRLNDLLKKNNEETGMNTP.IHVDAAAGGFAPFIYPELEWDFRPLVKSINVSGHK
 YGLVYAGTGVVYRAAEDLPEELI.FHINYHADQPTFTLNFSGKSSQIIAAYQIURL
 GFEGYKNVWNCIENWVLEKIEKTERFNIVSKQGVVFAVSLKDKDSFHFNEFISE
 MLRRFCWIVPAYMPADAQHI.TVLRVJREDFSRTLBRLLVADI.SKVLHELDLFLSKI
 SKMGIIEGTAENVKKEKWEKELMLVEVIVGWKRFVKRKKWNGVC"
 BASE COUNT 507 a 314 c 421 g 439 t
 ORIGIN

Query Match 57.7%; Score 870.2; DB 8; Length 1681;
 Best Local Similarity 75.1%; Pred. No. 5.7e-243;
 Matches 1110; Conservative 0; Mismatches 348; Indels 21; Gaps 1;
 QY 31 GAGGTCCTCGTCCACATCCACATTCGGATCACGTTACGTCCTCGTACCTTCACTTCTAGTTC 90
 Db 28 GATGAATCTGTGACCACTGTTCCGATCTCGTATGTTCCGACATACACTTCCCAAGTAT 87
 QY 91 AAGATCCCGAATACTCGATTCCTAAGGAGGGGTATCAGATCATCAACACGAGCTG 150
 Db 88 GAGATTGGTGAAGATTCGATACCGAAGACCGTGCATATCAGATCATAAAAGATGAGCTG 147
 QY 151 ATGCTTGCAGCGGAATCCACGTTGAACTTAGCTCTTGTGACGACATGATGGAGCCT 210
 Db 148 ATGCTTGTATGTTAAACCGAGGCTTAACCTAGCTTCTGTTGTGACTACATGGATGGAACCA 207
 QY 211 GAGTGTGAFAAATCACTCATGCTCCATCAACAGAACTATGTTGACATGACATGACAGATAC 270
 Db 208 GAGTGTGACAACTCATATGACTTATCAACAAAGACTAGCTTGTATGATGGATGAGTAC 267
 QY 271 CCGCTCACACCACTTCAGAACCGATGTTGAACTGTTGAACTGTTGCACTATTTCAATGCA 330
 Db 268 CCTGTCCAACTGAGCTCCAGAACCGATGTGAAACATTTATAGCTGTTCAATGGC 327
 QY 331 CCGTTAAGAGCGGAGACCCGCTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
 Db 328 CCACCTCGAGAACTTCAGACGCGGTGGAGTAGGACAGTGGTTCCTCAGAAGCCATC 387
 QY 391 ATGTTGGCCGTTGGCTTCAAGCGTAAATGGCGAACAAGCGCAAGCTGAAGCABA 450
 Db 388 ATGTTAGCCGGATGGCTTCAAAAGAAAATGGCAGAACAAACCGAAGCTGAGGGTAAA 447
 QY 451 CCGCTCGATTAACCCCAACTTGTCCGCGAGCAATGTTCAAGTGTGTTGGGAGAAATTC 510
 Db 448 CCTATGACAAACCCCAACTTGTCTGAGGCAATGTTCAAGTGTGTTGGGAGAAATTC 507
 QY 511 GCTAGTACTTTGAGGTTGAACTTAAGGAAGTGAATTCAGTGAAGGATACATGATGATG 570
 Db 508 GCTCCGTTACTTCGAGGTGAGCTAAAGGAAGTAAACCTAAGTGAAGTTACTACGTTAG 567
 QY 571 GACCTCAACAAGCTGTTGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 630
 Db 568 GATCCAGACAAAGACGAGAAATGGTAGACGAGAACACAATCTGTGTCCGACCCATATTG 627
 QY 631 GGTTCACCTTAAATGAGAAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCCGAA 690
 Db 628 GATCCACACTCAACCGTGAATTCGAGACGTGAACACGCTCAATGACTTGTCTAGTCAAG 687
 QY 691 AAGAACAAGAAACCGGATGGGATACCAATCCACGTTGGATCCGCAAGTGGAGGATTC 750
 Db 688 AAAAACGAGGAGACTGGTTGGAACAACACCGATCCACGTTGGATGACAGCAAGTGGAGGTT 747
 QY 751 ATTGCAACGTTTGTATCCGGAATGGAATGGACTTTAGACTTCCCTTGGTGAAGATG 810
 Db 748 ATAGTCCGTTTTATCTATCTGAAATAGAAATGGACTTTAGACTTCCCTTGGTGAAGAT 807
 QY 811 ATCAATGTCAGTGGTCAACAAGTATGACTTGTGTCAGGAGGATGTTGGTGTGATCTGG 870
 Db 808 ATCAAGTGAAGTGCACAAAGTATGGACTGGTCTATGCTGGTATTGGTGGTGGTGGTGG 867
 QY 871 AAGAAACAAGAGGATTTGCTGAGGAACCTCATCTCCATATCAATTAATCTTGGTCTGAC 930
 Db 868 AGGCCAGCAGAGGATTTGCTGAGGACTTATCTTCAATTAATTAATTAATCTTGGTCTGAT 927


```

331 CCCTTAGAGGCGGAGACCGCCGTCGGAGTAGGAAACCGCTTGGATCATCGGAGGCCATA 390
360 CCCTCAGGAAATCTAGACGGCGGTCGGAGTAGGAGACAGTTGGTCTTTCAGAAAGCCATC 419
391 ATGTTGGCCGCTTGGCCCTCAAGCCGTAATGGCAGAAACAAAGCCAAAGCTGAAGGCAAA 450
420 ATGTTAGCCGATTTGGCCCTCAAGAAATGGCAGAAACAAAGCCAAAGCTGAAGGTTAA 479
451 CCCGTCGATAAACCCAACTTGTCAACCGGAGCCAAATGTTCAAGFTGTTGGGAAATTC 510
480 CCCTATGACAAAACCCAACTTGTCAACCGGAGCCAAATGTTCAAGFTTCTGGGAAATTC 539
511 GCTAGTACTTTGAGTTGAATTAAGGAGTGAATTAAGGAGTGAATTAAGGAGTGAATTA 570
540 GCTCGGTACTTTCGAGTGGAGCTAAAGAAATTAAGGAGTGAATTAAGGAGTGAATTA 599
571 GACCCCTCAACAAGCTGTTGATGTTGATGAGAAACACCAATTTGTTGTCGGACATTC 630
600 GATCCAGACAAGCAGCAGAAATGTTAGACGAGAAACAAATCTGTGTCCAGCCATATTG 659
631 GGTTCACCTTAAATGAGAAATCGAAGATGTTAAACTTTGAAACGAATCTTTGGTCGAA 690
660 GGATCCACACTCAACGGTGTAGTCCGAAAGCTGAAAGCTCTCAATGACTTGTAGTCAAG 719
691 AAGAACAAAGAAACCGGATGGGATACACCAATCCACGTTGATGGGCAAGTGGAGGATTC 750
720 AAAAACGAGGAGACTGTTGGAAACACCCGATCCACGTTGATGGGCAAGTGGAGGATTC 779
751 ATTCACCCGTTTTGATTCGGAAATGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGT 810
780 ATAGCTCGTTTTACTATCTATCTGATTAAGTGGGACTTAGGCTTCTTGGTTAAGAGC 839
811 ATCAATGTAGTGTCAAGATGAGACTTGTGTACCGAGGAGTGGTGGGATCTGG 870
840 ATCAACGTGAGTGTCAAGATGAGACTTGTGTACCGAGGAGTGGTGGGATCTGG 899
871 AGAAACAAGAGGATTTGCCCTGAGAACTCATCTCCATATCAATTAATCTTGGTCTGAC 930
900 AGGCAGCAGAGGATTTACCTGAAAGCTTATCTTCAATTAATTAATCTTGGTCTGAT 959
931 CAACCCACTTTACTCTCAATTTCTCCAAAGGTTCAAAGTCAAAGTCAATGCTCAATACTAC 990
960 CAACCCACTTTACTCTCAATTTCTCCAAAGGTTCAAAGTCAAAGTCAATGCTCAATACTAC 1019
991 CAACCTTCCGATTTGGCCACGAGGTTACAGAAATGTGATGGAGAAATTCAGAGAGAA 1050
1020 CAGCTCATCTGCTTTGGATTCGAGGGGTACAAAATGTGATGGAGAAATTCAGAGAGAA 1079
1051 ATGATCGTCTAAGGAGGACTTTGAGAGACAGAAAGTTCACATCGTCTCAAAGGAC 1110
1080 ATGTTGTTCTCAAGAAAGTATAGAAHAAACAGAGGTTTTCAACATAGTCTCAAAGGAC 1139
1111 GAGGAGTGCCACTTGTGCTTTCTCCTTGAAGATAGCAGCTGTCACTGAGTTCGAA 1170
1140 CAAGGAGTGCAGTCTGCTCTCTCAAGGACCATAGTTTTCCCAACAGGTTCCGAG 1199
1171 ATCTCCGACATGCTTCGAGGATGGATAGTGGCCGCTTACAAATGCTCCAAAT 1230
1200 ATCTGTGATGCTACGTCGTTTTGGTGGATGCTCCAGCTTACATATGCTCCGCTG 1259
1231 GCACAACAATCATCTTCTGCTGTTGTTATCAGAGAAATTTCTCGAGAACTCGCT 1290
1260 GCACAGCACAATCAGGTTCTGCTGTTGTTGCTCATCAGGAAATTTCTCAAAGAACTCG 1319
1291 GAGAGACTTGTGATAGGAAATGATGCGTGGATGAGCTTCCATGAGCTTCTTCGAGA 1350
1320 GAGAGACTTGTGATAGGAAATGATGCGTGGATGAGCTTCCATGAGCTTCCATGAG 1379
1351 GTGATTCAAAATATCATCTGGCAAGAGAGAGTGAATCTAACAGCGGATTAATGATG 1410
1380 AT-----ATCTAAGAAATGGAATGAGGATCGCGAAAT 1418
1411 GTCAGGTTGAAGAGGAGGATATCGACAAGCAGAGAGATATCATCACTGCTGGGAGAG 1470

```

```

Db 1419 GTAAAGGAGAAAGATGGAAAGGAGATCTTGATGGAAGTTATTGTTGGATGGAGGAA 1478
QY 1471 TTTGTCCCGCAGCAGGAAAGAGAGTGGTATCTGCTAA 1509
Db 1479 TTTGTGAAGGAGGAAAGATGAATGTTGTGCTAA 1517

```

Search completed: October 22, 2003, 12:44:29
Job time : 5522 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: October 22, 2003, 10:16:42 ; Search time 2998 Seconds
(without alignments)
12233.307 Million cell updates/sec

title: US-10-006-852-1
erfect score: 1509
equence: 1 atggtctctcccacgcgt.....agcagtggtatctgctaa 1509

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

searched: 22781392 seqs, 12152238056 residues

total number of hits satisfying chosen parameters: 45562784

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

database :

- 1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estma:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: gb_gss1:*
28: gb_gss2:*
29: gb_gss3:*

Table with 5 columns: ID, Score, Length, DB ID, Description. Contains entries for Zea mays and other species.

ALIGNMENTS

RESULT 1
AY103733
LOCUS AY103733 1168 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0153431 mRNA sequence.
ACCESSION AY103733
VERSION AY103733.1 GI:21206811
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1168)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whittitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1168)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with 4 columns: result No., Score, Match Length, DB ID, Description. Lists search results for Zea mays.

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from Zmdb: www.zmdb.iastate.edu.

Location/Qualifiers
1..1168

/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:638900"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project."

BASE COUNT 275 a 332 c 345 g 216 t

Query Match 40.0%; Score 603.2; DB 11; Length 1168;
Best Local Similarity 74.2%; Pred. No. 1.9e-165;
Matches 775; Conservative 0; Mismatches 268; Indels 1; Gaps 1;

6 GCTCTCCACGCGGTATCGGATCGGAGTCTCCGTCACCTCCACATTCGCATCACGTTA 65
118 GCACGCGAGCTCCGCGCGGACGACGACGCGGTGCGCTCCACCTTCGCGTCCGCGTA 177
66 CGTCCGTTACTTCCATTCCTAGTCTCAAGATCGGAAAACCTCGATTCCTAAGGAGGGCC 125
178 CGTGGCGAGCGCTCCCGCGGTACCGGATCGCGAGCGGTCGATCCCGCGGGAGCGCC 237
126 GTATCAGATCATCAACGAGCTGATGCTTGACCGGAAATCCACGGTTGAACTTAGCGTC 185
238 GCAGCAGATCATCAGCAGAGCTCATGCTGACGCGCAACCGCGCTGAACCTGGCGTC 297
186 CTTTGTGAGCAGATGATGAGCGCTGAGTGTGATAACTCATCATGCTCCTCCACACAA 245
298 CTTGCTFCAACAGTGGAGCGCGGATGACAAAGCTCATCGACTCCATCAACA 357
246 GAACTATGTTGATCAGCAGTACCCCGTACCACCGAACTTCAAGAACCCGATGTTGAA 305
358 GAACATCGTGCATGAGAGTACCCCGTACCACCTGAGCTCCAGAACCCGTCGTTAA 417
306 CATGATGACATCATTTCAATGCAACCGTTAGAGAGGAGACCGCGCTCGAGATAGG 365
418 CATGATGATCACCCTGTTCAACCGCGCATCAAGGAGGACGACAGCTATTGGGGTCGG 477
366 AACCGTTGATCATCGGAGCCATATGTTGGCGGTTGGCTTCAAGCGTTAAATGGCA 425
478 AACATGGGATCTCGAAGCGATCATGCTTCAGGTCTGGGTTCAAGAGAAATGGCA 537
426 GAACAAGCGCAAGGTGAGGCAAAACCCGTCGATAAACCACCAATFTFCACCGGAGCAA 485
538 GAACAAGAGAAAGGACGAGGGAAAACCGTGTGACAGACCCCAACATCGTCAACCGTGC 597
486 TGTTCAGTGTCTGGGAGAAATTCGTTAGTACTTTGAGGTTGAACTTAAGGAGTAA 545
598 CGTTCAGGTTCTGGGAGAAATTTGCAAGTATTTTGGTGGGACTGAAGAGGTTAA 657
546 ATTGATGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605
658 GCTATCAGAAAGGTTACTAGTTCATGACCCCGTCAAGGCTGTGAGATGGTGGACGAA 717
606 CACATTTGTTGGGCAATTCCTTGTTCCTTAAATGAGAAATTCGAAAGTGTAA 665
718 CACTATCTGTGCGCGCATCTTGGGTCGACTCTCCCGGAGATTCGAGGATGTCAA 777
666 ACTCTTGAACGATCTCTTGGTGGAAAGCAAGAAACCGGATGAGTACACCAATCCA 725
778 GCAATTTGAACGATCTCTTACAGAAAGCAAGAAACTGGTGGGACGTCCTTATCCA 837
726 CGTGGATGGCAAGTGGAGGATTCATTCACCCGTTTTTGTATCCCGGAAATGGAATGGGA 785

838 CGTCGACCGCGGAGCGGAGGTTTCATAGCGCTTTCCTACCTGAGCTGAGTGGGA 897
786 CTTTACACTTCCTTCGTCGAGAGTATCAATCTGATGATGTCACAGATGAGTGGTGTGTA 845
898 CTTACAGCTGCTCTGTTGAAGAGCATCAATCTAGCGGGCAAGTACGGCTCGTCTA 957
846 CGCAGGATTTGTTGGTGTGATCTGGGAAACAAAGAGGATTTGCTGAGGAACATCATCTT 905
958 CCCTGGTGTGGCTGGTCTATCTGGCG-GACAAGGATGATGCTGAGGAGCTCATCTT 1016
906 CCATATCAATATCTTGGTGTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTTC 965
1017 CCATATCAATATCTTGGGAGCCAGCCAGCTACGTTACGCTCAACTTCTCCAAAGGTTTC 1076
966 AAGTCAAGTCAATGCTCAATCTTACCACACTTATCCGATTCGATGGCCACGAGGTTACAGAAA 1025
1077 CTTGTCAGATCTTGGCAGTACTATCACTCACTTTCGCTTAGGCTTGGAGGGTACAGAAA 1136
1026 TGTGATGGGAATTCAGAGAAA 1049
1137 CATCATGGAACTGCCAGAGAAA 1160

RESULT 2
AV541881
LOCUS AV541881 Arabidopsis thaliana roots Columbia linear EST 07-SEP-2000
DEFINITION cDNA clone RZ172907F 3', mRNA sequence.

ACCESSION AV541881
VERSION AV541881.1 GI:8703640
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsid.

REFERENCE 1 (bases 1 to 570)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 19,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..570
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZ172907F"
/tissue_type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 161 a 104 c 156 g 149 t
ORIGIN

Query Match 37.7%; Score 568.4; DB 9; Length 570;
Best Local Similarity 99.8%; Pred. No. 2e-155;
Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 351 CGCCGTCGAGTAGGAAACCGTTGGATCATCGGAGCCATAATGTTGGCGGTTGGCCTT 410
Db 1 CGCCGTCGAGTAGGAAACCGTTGGATCATCGGAGCCATAATGTTGGCGGTTGGCCTT 60
QY 411 CAAGCGTAAATGGCAGAACAGCAAGCTGAAGGCAACCCGTCGATAAACCCCAACAT 470

61 CAAGCGTAATGGCAGAAACAAAGCGCAAAAGCTGAAGCGCAAAACCCGTCGATAAAACCCCAACAT 120
 471 TGTACCCGAGCCAAATGTTCAAGTGTGTTGGAGAAATTCGCTAGGTACTTTGAGGTTGA 530
 121 TGTACCCGAGCCAAATGTTCAAGTGTGTTGGAGAAATTCGCTAGGTACTTTGAGGTTGA 180
 531 ACTTAAGGAAGTGAATAATGAGTGAAGGATACTATGTGATGACCCCTCAACAAGCTGTTGA 590
 181 ACTTAAGGAAGTGAATAATGAGTGAAGGATACTATGTGATGACCCCTCAACAAGCTGTTGA 240
 591 TATGGTTCATGAGAACACCAATTTGTTGCGGACATCTTGGTCCACTCTTAATGAGAGA 650
 241 TATGGTTCATGAGAACACCAATTTGTTGCGGACATCTTGGTCCACTCTTAATGAGAGA 300
 651 ATTCCGAAGTGAATAATGAGTGAAGGATACTATGTGATGACCCCTCAACAAGCTGTTGA 710
 301 ATTCCGAAGTGAATAATGAGTGAAGGATACTATGTGATGACCCCTCAACAAGCTGTTGA 360
 711 GGATACACCAATCCACGFGAGTGGCGCAAGTGGAGGATTCATTCACCGCTTTTGTATCC 770
 361 GGATACACCAATCCACGFGAGTGGCGCAAGTGGAGGATTCATTCACCGCTTTTGTATCC 420
 771 GGAATTCGAAGTGGAGTTCAGACTTCCCTTGGTGAAGGATTCATTCAGTGGTTCACAA 830
 421 GGAATTCGAAGTGGAGTTCAGACTTCCCTTGGTGAAGGATTCATTCAGTGGTTCACAA 480
 831 GTATGGACTTGTACCGAGGATTTGGTGGTGTATCTGGAGAAACAAAGAGGATTTGCC 890
 481 GTATGGACTTGTACCGAGGATTTGGTGGTGTATCTGGAGAAACAAAGAGGATTTGCC 540
 891 TGAGGAATCTATCTCCATATCAATTAATCT 920
 541 TGAGGAATCTATCTCCATATCAATTAATCT 570

SULT 3
 554189
 CUS AV554189 548 bp mRNA linear EST 06-SEP-2000
 FINITION AV554189 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
 CESSION AV554189 cDNA clone R28306R 5', mRNA sequence.
 RSION AV554189.1 GI:8725602
 YWORDS EST.
 ORGANISM Arabidopsis thaliana (chale cress)
 URCE Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 548)
 Asamizu, E., Makamura, Y., Sato, S. and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)
 MEDLINE 20363093
 PUBMED 10307847
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
 Location/Qualifiers
 1. 548
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strains="Columbia"
 /db_xref="taxon:3702"
 /clone="R28306R"
 /tissue_type="roots"
 /clone_lib="Arabidopsis thaliana roots Columbia"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"
 ATURES source

BASE COUNT 164 a 114 c 132 g 138 t
 ORIGIN
 Query Match 36.3%; Score 548; DB 9; Length 548;
 Best Local Similarity 100.0%; Pred. No. 2e-149;
 Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 849 AGGGATTGGTTGGTGGATCTGGAGAAACAAAGAGGATTTGCCCTGAGAACTCATCTTCCA 908
 Db 1 AGGGATTGGTTGGTGGATCTGGAGAAACAAAGAGGATTTGCCCTGAGAACTCATCTTCCA 60
 QY 909 TATCAATTAATCTTGGTGGTGCACCAACCCACCTTTACTCAATTTCTCCAAAGGTTCAAG 968
 Db 61 TATCAATTAATCTTGGTGGTGCACCAACCCACCTTTACTCAATTTCTCCAAAGGTTCAAG 120
 QY 969 TCAAGTCAATTTGGTTCATCTACCACTTATCCGATTCGANTGGCCACGAGGTTACGAAATGT 1028
 Db 121 TCAAGTCAATTTGGTTCATCTACCACTTATCCGATTCGANTGGCCACGAGGTTACGAAATGT 180
 QY 1029 GATGGAGAATTCGAGAGAGATATGATCGTCTTAAGGGAAAGACTTGAGAAACACAGAAAG 1088
 Db 181 GATGGAGAATTCGAGAGAGATATGATCGTCTTAAGGGAAAGACTTGAGAAACACAGAAAG 240
 QY 1089 GTTCAACATCGTCTCAAGAGAGAGGAGTGCACCTTTCGCTTCTCCTTCAAGATAG 1148
 Db 241 GTTCAACATCGTCTCAAGAGAGAGGAGTGCACCTTTCGCTTCTCCTTCAAGATAG 300
 QY 1149 CAGCTGTACACTGAGTTCGAAATCTCCGACATGTTCCAGGATGATGGATAGTGGC 1208
 Db 301 CAGCTGTACACTGAGTTCGAAATCTCCGACATGTTCCAGGATGATGGATAGTGGC 360
 QY 1209 GGCCTACAAATCGCTCCAAATGCAACAACATCACTGTTTCGTTGGTTATCAGAGA 1268
 Db 361 GGCCTACAAATCGCTCCAAATGCAACAACATCACTGTTTCGTTGGTTATCAGAGA 420
 QY 1269 AGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAAAGTGCATCGTGA 1328
 Db 421 AGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAAAGTGCATCGTGA 480
 QY 1329 GCTCGATGAGTTCCTTCGAGAGTGCATTCACAAAATATCATTGGACAGAGAGAGTGA 1388
 Db 481 GCTCGATGAGTTCCTTCGAGAGTGCATTCACAAAATATCATTGGACAGAGAGAGTGA 540
 QY 1389 ATCTAACA 1396
 Db 541 ATCTAACA 548
 RESULT 4
 CB655721
 LOCUS OSJNEC09G09.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
 DEFINITION clone OSJNEC09G09 5', mRNA sequence.
 ACCESSION CB655721
 VERSION
 KEYWORDS EST.
 SOURCE GI:296659446
 ORGANISM Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 811)
 Jantanasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
 Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished
 Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967

Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aaa cga cgg cca gtg
 BACKWARD: gga aac agc tat gac cat g
 Plate: 09 row: G column: 09
 Seq primer: gta aaa cga cgg cca gtg.
 Location/Qualifiers
 1..811
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNC09609"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNC"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
 XhoI 182 a 236 c 244 g 149 t

Db 661 TCGTCTACCCGGATCGGGTGTGCACTTGAGGAGCAAGGAGGATCTGCCTGAGGAC 720
 QY 899 TCATCTCCATATCAATATATCTTGGTGTGACCAACCCACCTTTACTCTCAATTTTCCA 958
 Db 721 TCATCTCCATATCAATACCTCGGGCGGACGACCCACCTTCACTCAACTTCTCCA 780
 QY 959 AAGGTTCAAGTCAAGTCAATGCTCAATCTA 989
 Db 781 AGGGTTACAGCCAGGTCATTGCACAGTATTA 811

RESULT 5
 BF275438
 LOCUS
 DEFINITION
 Gossypium arboreum 7-10 dpa fiber library Gossypium
 arboreum cDNA clone GA_Eb0024A19f, mRNA sequence.
 ACCESSION
 BF275438
 VERSION
 BF275438.1 GI:11206508
 KEYWORDS
 EST.
 SOURCE
 Gossypium arboreum
 ORGANISM
 Gossypium arboreum

BF275438 837 bp mRNA linear EST 07-MAR-2001
 GA_Eb0024A19f Gossypium arboreum 7-10 dpa fiber library Gossypium
 arboreum cDNA clone GA_Eb0024A19f, mRNA sequence.
 BF275438
 BF275438.1 GI:11206508
 EST.
 Gossypium arboreum
 Gossypium arboreum
 Gossypium arboreum
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 REFERENCE
 1 (bases 1 to 837)
 AUTHORS
 Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
 ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
 TITLE
 An integrated analysis of the genetics, development, and evolution
 of the cotton fiber
 JOURNAL
 Unpublished
 COMMENT
 Contact:Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAAATGACTCATTATAGGG
 High quality sequence start: 2
 High quality sequence stop: 718.

SE COUNT 182 a 236 c 244 g 149 t
 IGIN

Query Match 34.9%; Score 526.2; DB 14; Length 811;
 Best Local Similarity 78.1%; Pred. No. 6.3e-143;
 Matches 633; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

FEATURES
 source

1..837
 /organism="Gossypium arboreum"
 /mol_type="mRNA"
 /strain="AKA"
 /cultivar="8400"
 /db_xref="taxon:29729"
 /clone="GA_Eb0024A19f"
 /tissue_type="Fibers isolated from bolls harvested 7-10
 dpa"
 /lab_host="E. coli"
 /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
 /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
 244 a 171 c 198 g 223 t 1 others

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

Query Match 34.8%; Score 525.2; DB 10; Length 837;
 Best Local Similarity 78.9%; Pred. No. 1.3e-142;
 Matches 637; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

QY 1 ATGGTCTCTCCACGCCGTATCGGAGTCGGAGTCGCGTCCGTCACATCCGATCA 60
 Db 32 ATGGGTATATCAAGACAGCATCTGAAACTGATGTCTCAGTTCATCCACTTTCCCTCT 91
 QY 61 CGTACGTCGGTACTTCACTTCTAGGTTCAAGATCGCGGAAACTCGATTCCTAAGGAA 120
 Db 92 CGTATGTCGAACTCACTGCCAGTTCAAAATGCCGAAACTTCCATCCAAAGAG 151
 QY 121 GCGCGTATCAGATCATCAACGACGAGTGTGTTGACGGGAATCCACGGTTGAACTTA 180
 Db 152 GCTGCTTATCAGATCATCAATGATGAACCTGATGCTTTGATGGCAACCCCAAGGTTGAACCTT 211
 QY 181 GCCTCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAATCATCATGTCTCCCTC 240

179 TAGCCTCTTTGTGACGACATGATGGAGTGGAGTGTGATGATAAATCACTCATATGCTCTCCA 238
 1 TCGCGTCTGTTCCACACGATGGATGGAGCCGAGTGCAGCAAGCTCATCCAGGCTCCG 60
 239 TCAACAGAACTATGTTGATGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 298
 61 TCAACAGAACTATGTTGATGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 120
 299 GTGTGAACATGATTGCACATCTAATCAATGACCGTTAGAAAGGCGGAGCGGAGCGGCTGC 358
 121 GTGTGAACATGATTGCACATCTAATCAATGACCGTTAGAAAGGCGGAGCGGAGCGGCTGC 180
 359 GAGPAGAACCGTTGGATFCAGGAGCCCAATGTTGGCCGTTGGCTTCAAGCGTA 418
 181 GAGTCCGCACTGTCCGCTGCTGAGGCCATCATGCTCCGCGGTTTGGCTTCAAGAGGA 240
 419 AATGCCAGAACAGCCGAAAGTGAAGGCAACCCGTCGATTAACCAACCAATGTCACCG 478
 241 GGTGGCAGAACAGATGAAGCAGCCGCGCAAGCCATGCCAAGCCCTAACATTTGTCACCG 300
 479 GAGCAATGTTCAAGTGTGTTGGAGAAATTCGCTAGTGTACTTTGAGGTTGAACTTAAGG 538
 301 GCGCAATGTTCAAGTGTGTTGGAGAAATTCGCTAGTGTACTTTGAGGTTGAACTTAAGG 360
 539 AAGTGAATGAGTGAAGGATCATGATGATGAGCCCTCAACAGCTGTTGATATGTTG 598
 361 AAGTGAAGCTGAGTGAAGGATCATGATGATGAGCCCTCAACAGCTGTTGATATGTTG 420
 599 ATGAGAACCAATTTGTTGCGGACATCTTGTGTTTCCACTCTTAATGGAAATTCGAAG 658
 421 ACGAGAACCAATTTGTTGCGGACATCTTGTGTTTCCACTCTTAATGGAAATTCGAAG 480
 659 ATGTTAACTTCAACGATCTCTTGTGCAAAAAGAAACCAAGAAACCGGATGGATACAC 718
 481 ACGTGAAGCTGCTCAACGATCTGCTACCAAGAAAGACGCTGAACAGGCTGGGACAGC 540
 719 CAATCCAGTGGATTCGGCAAGTGGAGGATTCATTTGACCCGTTTTTTTATCCGGAATGG 778
 541 CGATCCAGTGGACGCGGAGCGGGTTTCATCGCGGCTTCTGTAACCCGAGGCTGG 600
 779 AATGGACTTATGACTTCCCTTGTGAAGAGTATCAATGATGAGTGGTCACAAGTATGGAC 838
 601 AGTGGACTTCCCGCTCCCGTGTGAGAGCATCAACGTTGAGCGGCAAGTATCGGCC 660
 839 TTGTGTACGCGGATGTTGTTGGTGTATCGAGAAAACAAGAGGATTTCCCTGAGGAA 898

722 TCCACGTGGATCCGCGCAAGTGGAGGATTCATTGACACCGTCTTTTGTATCCCGAAATTTGGAAT 781
 751 TTCAATGGATGACAGTGGTGGTTCATTGCTCCATTTTATCCAGAACTTGGAT 810
 782 GGGACTT 788
 811 GGGATTT 817

SULT 7
 040860
 CUS
 BE040860 821 bp mRNA linear EST 07-JUN-2000
 OF12G05 of *Oryza sativa* cDNA 5', similar to glutamate decarboxylase
 1, mRNA sequence.
 BE040860
 BE040860.1 GI:8336367
 EST.
 Oryza sativa
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 821)
 Bohnert,H.J.; Borchert,C.; Brazille,S.; Brooks,J.; Eaton,M.; Ferreira,
 H.; Kawasaki,S.; McColough,A.; Michalowski,C.B.; Palacio,C.;
 Scara,G.; Wheeler,M. and Zepeda,G.R.
 Functional Genomics of Plant Stress Tolerance
 Unpublished
 Contact: Michalowski,C.B.
 University of Arizona
 Bio Sciences West room 513, Tucson, AZ 85721, USA
 Tel: 520-621-7982
 Fax: 520-621-1697
 Email: cbm@u.arizona.edu
 An open reading frame exists.
 Insert Length: 1 Std Error: 0.00.

Location/Qualifiers
 1. 821
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /strain="Pokkali"
 /db_xref="taxon:4530"
 /tissue_type="entire plant"
 /dev_stages="2 weeks"
 /clone_lib="OP"
 /note="1 week 150mM NaCl"
 185 a 231 c 254 g 149 t 2 others

SE COUNT
 IGIN
 Query Match 33.7%; Score 509.2; DB 10; Length 821;
 Best Local Similarity 78.0%; Pred. No. 6.3e-138;
 Matches 613; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

132 GATCATCAACGACGAGTGCCTTGGCGGAATCCAGGTTGAACTAGCCTCTTTGT 191
 5 GAGGATCAACGACGAGTGCCTTGGCGGAATCCAGGTTGAACTAGCCTCTTTGT 64

192 GACGACATGGATGGAGCCTGTGATGATPAACTCATCATGTCCTCCATCAACAAGAACTA 251
 65 CACCACGTTGGATGGACCGGATGGGACAGCTCATCAGGCCTCCGTCACAAAGAACTA 124
 252 TGTTCAGATGACGAGTACCCCGTCCACCAGAACTTCCAGAACCGATGTGTGAACATGAT 311
 125 CGTCGACATGGACGAGTACCCCGTCCACCAGGCTCCAGAACCGATGTGTGAACATGAT 184

312 TGCACATATTTCATGACCGTTAGAGAGCGGAGACCCCGTCCGAGTAGGACCGT 371
 185 TGCACACCTTTCAATGCTCTCTAGGGGACTCTGAAACCGCGGTCGGGACCTGT 244
 372 TGGATCATCGGAGGCCATAATGTTGGCGGTTGGCCCTTCAAGCGTAAATGGCAGAA 431
 245 CGGCTCGTCTGAGGCCATCATGCTCCCGGTTTGGCCCTTCAAGAGGAGGTGGCAGAA 304

432 GCGCAAGCTGAAGCGAAACCCCGTCCGATAAACCACATTTGTCACCGGAGCCAAATGTTCA 491
 305 GATGAAGGCGAGCCGCAAGCCATGCGACAGCCTAACATTTGTCACCGGCGCAATGTTCA 364
 492 AGTGTGTTGGGAGAAATTCGCTAGGTAATTTGAGTTGAACTTAAAGGAAGTGAATTTGAG 551
 365 AGTTTGTGGGAGAAATTCGCGCGGATCTTCGAGGTTGAGCTCAAGGAAGTGAAGCTGAG 424
 552 TGAAGGATACTATGTGATGGACCCCTCAACAAGCTCTTGATATGTTGATGAGAACACCAT 611
 425 TGACGGCTACTACGTCATGGACCCAGCTAAGGCCCTGGATATGTTGTCGACGAAACACCAT 484
 612 TTTGTTGGGACATTTCTTTGGTTCCACTCTTAATGGAGAAATTCGAAAGATGTTAAACTCTT 671
 485 CTGCGTCCGGCGGATCTTCGCGTTCGACGCTGAACGGGAGTTTCGAGGACCTGAAGCTGCT 544
 672 GAACGATCTCTTGGTGGAAAAGAAACAAGAAACCCGATGGGATGACCAATCCACGTTGA 731
 545 CAACGATCTCTCACCAGAGAAAGCCTGAAACAGGCTGGGACACCCGATCCACGTTGA 604
 732 TCGCGCAAGTGGAGGATTCATTGTCACCGCTTTTGTATCCGAAATTTGAAATGGACTTTAG 791
 605 CGCGGCGAGCGGGTTCATCGCCGTTCTCTGACCCGAGCTGGAGTGGGACTTTTCG 664
 792 ACTTCCCTTGGTGAAGATTCATGATGAGTGGTCAACAGTATGAGCTTGTGTACCCGAGG 851
 665 GCTCGCGTGTGAAAGATCAACGTTGAGCGCGGACCAAGTACCGGCTCGTCTACGCGCG 724
 852 GATTGTTGGTGTGATCTGGAGAAACAAGAGATTTGCTGAGGAACTCATCTTCCATAT 911
 725 GATTGTTGGTGTGATCTGGAGAAACAAGAGATTTGCTGAGGAACTCATCTTCCATAT 784

912 CAATTA 917
 785 AACTTA 790

AV552358 554 bp mRNA linear EST 06-SEP-2000
 AV552358 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
 cDNA clone RZ27f07R 5', mRNA sequence.
 AV552358
 AV552358.1 GI:8723771
 EST.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 554)
 Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)
 20363093
 10907847
 PUBMED
 Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

Location/Qualifiers
 1. 554
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_lib="RZ27f07R"
 /tissue_type="roots"
 /clone_lib="Arabidopsis thaliana roots Columbia"

RESULT 8
 AV552358
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 FEATURES
 source

961 GGTTCAGTCAAGTCAATGCTCAATCACTACCACTTATCCGATTTGGCCACGAGGGTTAC 1020
 |||||
 601 GGTTCAGTCAAAATAATGCTCAATCACTATCACTGTTTCCGAGGATAC 660
 |||||
 1021 AGAAATGTGATGGAGATTCGAGAGAGAAATATGATCGTCTTAAGGGAAGCATTGAGAG 1080
 |||||
 661 AAGAGTATCATGGAAACTGNCAGAAAATGCAAAATGTTCTAAAGAAAGGTTTGGAGAAA 720
 |||||
 1081 ACAGAAAGGTTCAACATCGTCTCAAGGACGAGGGAGTGCACCTTFCGCTTCT 1135
 |||||
 721 ACCGACGGTTCACATCATCTCTAAAGAACACGCGCTCCACCTTGTAGCCTTCT 775
 |||||

CA813481 786 bp mRNA linear EST 11-APR-2003
 CA48LN101F-E8 Cabernet Sauvignon Leaf - CA48LN Vitis vinifera cDNA
 clone CA48LN101F-E8 5', mRNA sequence.

CA813481
 CA813481.1 GI:26262418
 EST.
 Vitis vinifera
 Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids
 ; Vitaceae; Vitis.
 1 (bases 1 to 786)
 Goes da Silva, F., Lim, H., Iandolo, A., Baek, J., Jones, K., Walker
 , M.A. and Cook, D.R.
 Transcriptional responses of Vitis vinifera to infection by the
 bacterial pathogen Xylella fastidiosa
 Unpublished
 Contact: Doug Cook
 CES Genome Facility
 UC Davis Department of Plant Pathology
 1 Shields Ave., Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcook@ucdavis.edu
 Seg primer: GTTATCAGTCGACGGTACC.

source
 1. 786
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CA48LN101F-E8"
 /sex="hermaphrodite"
 /dev_stage="late season sample"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Leaf - CA48LN"
 /note="Organ: Leaf; Vector: pDNR; Site: SflI; Site.2:
 SflI; CA48LN is a cDNA library of Cabernet Sauvignon
 leaves. The leaves were collected on September 20, 2001,
 in Napa Valley, California, and represent leaves in late
 season development. These leaves were asymptomatic and
 verified to be non-infected with the bacterial pathogen,
 Xylella fastidiosa, based on a diagnostic assay using PCR
 and Xylella-specific primer pairs. cDNAs were made by
 oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGAGTGTCCATTCAGCGCGG-3', and
 5'-ATTCTAGAGCGGCGCGGCGGACATG-dT(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."
 SE COUNT 222 a 163 c 200 g 201 t
 IGIN
 Query Match 32.9%; Score 496.6; DB 14; Length 786;
 Best Local Similarity 77.8%; Pred. No. 3.1e-134;
 Matches 612; Conservative 0; Mismatches 174; Indels 1; Gaps 1;
 523 GAGGTTGAACCTTAGGAAGTGAATTTGATGAGGAGTACTATGATGGAGCCCTCAACAA 582

Db |||||
 1 GAGTGGAAATTAAGCAAGTGAAGCTGACGATGCTACTATGTGATGAGCCCTGTCAAA 60
 |||||
 QY 583 GCTGTGATATGTTGATGAGAAACACCACTTTGTTGGGACATTTCTTTGGTTCCACTCTT 642
 |||||
 Db 61 GCAGTGGAGATGGTGGACGAGAACACCACTTGTGTGCTGCTAICCTTGGTTTACCCCTC 120
 |||||
 QY 643 AATGGAGAAATTCGAAGATGTTAAACTTTGAACGATCTTTTGGTCGAAAAAGAACAAAGAA 702
 |||||
 Db 121 AATGGAGAAATTCGAAGATGTTAAACACTTTGAATGATCTTTTGCACAAAGAAAAACAAGAA 180
 |||||
 QY 703 ACCGATGGGATACACCAATCCAGTGGATGGGCAAGTGGAGSATTTCATTGCACCGTTT 762
 |||||
 Db 181 ACAGGATGGGATACCAATCCATGTAGATGAGCAAGTGGTGGCTTTTATTGCACCCCTT 240
 |||||
 QY 763 TTGTATCCGGAAATTCGAATGGGACTTTAGACTTCCCTTTGGTGAAGAGTATCAAATGTGAGT 822
 |||||
 Db 241 CTTACCCAGAGCTGGAGTGGGACTTCCGCTTCCACTTGTGAAAAGCATTAACGTTAGC 300
 |||||
 QY 823 GGTCAAGATATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 882
 |||||
 Db 301 GGCCACAGTACGGGCTTGTGTATGCTGGCATAGTTGGTGTGATTTGGAGGAGTAAAGAG 360
 |||||
 QY 883 GATTTGCCTGAGAACTCATCTTCCATATCAATTAATCTTTGGTGTGACCAACCCACCTTT 942
 |||||
 Db 361 GACTTCCCGGATGAGCTCATCTTTTCAATATCAACTATCTTTGGAGCCGCAACCTTACCTTC 420
 |||||
 QY 943 ACTTCAATTTCTCCAAAAGTTCAAGTCAAGTCAATGCTCAATTAATCACTCAACTTATCCGA 1002
 |||||
 Db 421 ACCCTCAACTTCTCCAAAAGTTCTAGTCAAGTATTTGCTCAGTATTAATCAACTCAFTTGC 480
 |||||
 QY 1003 TTGGCCACGAGGTTTACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1062
 |||||
 Db 481 TTGGCCCACTAGCGGATATAGGAACTATTTGGAAAATCTGCAAGAAAATTCGAAGGTTGCTC 540
 |||||
 QY 1063 AGGGAAGGACTTTGAGAAGACAGAAAAGTTCAACATCGTCTCAAAGGACGAGGAGTGGCA 1122
 |||||
 Db 541 AAGAGAGGATTTGGAGAAGACAGGGGACTTCAACATTTGTTCTAAAGAAATGGAGTGGCA 600
 |||||
 QY 1123 CTTGTGCTTTCTCCCTTGAAGATAGCAGCTGTCTACACTGATTTGAAATCTCCGACATG 1182
 |||||
 Db 601 TTAGTAGCTTCTCTTTGAAAGACACAGCTGTCAACAATGATGATGATGATGATGATGATGATGAT 660
 |||||
 QY 1183 CTTGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1242
 |||||
 Db 661 CTGCCAGCTATGGATGGATCGTGGCAGCTTATACCATACCGCCGGATGCACAACTGTG 720
 |||||
 QY 1243 ACTGTTCTTCGTTGTTATCAGAGAAGATTTCTCGAAACTCGCTGAGAGACTTGTG 1302
 |||||
 Db 721 ACTGTTCTTCGTTGTTATCAGAGAAGATTTCTCGAAACTCGCTGAGAGACTTGTG 779
 |||||
 QY 1303 ATCGATA 1309
 |||||
 Db 780 ACAGATA 786
 |||||
 RESULT 11
 BQ971999
 LOCUS BQ971999
 DEFINITION Helianthus annuus sunflower RHA801 Helianthus annuus cDNA
 clone OHB9G09, mRNA sequence.
 ACCESSION BQ971999
 VERSION BQ971999.1 GI:22389520
 KEYWORDS EST.
 SOURCE Helianthus annuus (common sunflower)
 ORGANISM Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asterales; Asteraceae; Asteroidae;
 Heliantheae; Helianthus.
 REFERENCE 1 (bases 1 to 762)
 AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevaller, P., Ziegler, J., Ellison

uninoculated leaves were harvested 20 hr post-inoculation (Wei, Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one cDNA library was made, and 1 million pfu were in vivo excised to give pluescript SK(-) cDNA phagemids (Choi, Close). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Cates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders>. Also see Close TJ, Wing R, Kleinmors A, Wise R (2001). Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/gppages/bgm/31/cover.html>)"

209 a 187 c 249 g 175 t 3 others
 .IGIN
 Query Match 32.1%; Score 484.8; DB 10; Length 823;
 Best Local Similarity 77.2%; Pred. No. 9.2e-131;
 Matches 602; Conservative 0; Mismatches 177; Indels 1; Gaps 1;

100 GAAAATCCACGGTGAACCTAGCTCCCTTTGTGACGACATGATGAGCCCTGAGTGTGAT 219
 43 GAAACTCGATCCCAAGAGCGCGGATCAGATCAACAGCAGAGCTGATGCTGAC 102
 160 GGGAAATCCACGGTGAACCTAGCTCCCTTTGTGACGACATGATGAGCCCTGAGTGTGAT 219
 103 GGCACCCCGCGCTGAACCTGCGCTCCCTTGTGACGACATGATGAGCCCTGAGTGTGAT 219
 220 AAACATCATGTCCTCCATCAGCAAGAACTATGTTGATGATGAGCCCTGAGTGTGAT 219
 163 AAGCTCATATGATGCTCCGTCACCAAGAACTACGTCGATGAGCCCTGAGTGTGAT 219
 280 ACCGAATTCAGAACCGATGTTGTAACATGATGACATATTCATGACCCGCTGAG 339
 223 ACCGAGTCCAGAACCGTGTGTAACATGATGATGATGATGATGATGATGATGAT 282
 340 GAGCGGAGACCGCGTGGAGTAGAACCGTGTGATGATGATGATGATGATGATGAT 339
 283 GAGGACGACACAGCTATCGGAGTCTCGACGCTGCGGCTTTCGGAAGCAATCATGCTTGA 342
 400 GGTGTCCTTCAAGCGTAAATGGCAGACACAGCCGAAGCTGAGGCAACCCGCTGAT 459
 343 GGCCTGGCTTCAAGAGAAATGGCAGACACAAATGAAAGAGCAGGGAGCCATGCGAC 402
 460 AAACCCAACTTGTCCCGGACCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGTAC 519
 403 AAACCTAACTTGTACTGTTGTTCAATGTTCAAGTGTGTTGCTGGAGAAATTTGCTAGTAT 462
 520 TTTGAGTTGAACTTAAGGAAGTAAATGATGATGATGATGATGATGATGATGATGAT 579
 463 TTTGAGTTGAACTTAAGGAAGTAAATGATGATGATGATGATGATGATGATGATGAT 522
 580 CAACTGTTGATATGTTGATGAGAACCACTTGTGTTGCGGACATCTTGGTTCCTACT 639
 523 AAGCTGTTGAAATGTTGATGAGAACCACTATATGTTGCGCCGCACTCTGGATCTACT 582
 640 CTTAATGGAGAAATTCGAAGATGTTAACTTCTTGAACGATCTCTTGGTGAAGAAACAAA 699
 583 CTCATGGAGAGTACGAGATGCAAACTGTTGAATGACCTTCTTGTGAGAGAAACAAA 642
 700 GAAACCGGATGGATACACCAATCCACGTGATGATGATGATGATGATGATGATGATGAT 759
 643 GAAACAGGTTGAAACGTCGCCGATTCATGTTGATGATGATGATGATGATGATGATGAT 702

QY 760 TTTTTCATCCGAAATGGAATGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTG 819
 |||||
 Db 703 TTTTTCATCCGAGCTTGAATGGAGCTTCAAGCTACCATTTGTTGAAGAGCATCAACGGT 762
 |||||
 QY 820 AGTGGTCAAGTA-TGGACTTGTGTACCGAGGATGGTGGGTGATCTGGAGAAACA 878
 |||||
 Db 763 AGGGGCAACAAGTATTGCCCTTGTGTACCCCTGGGTTGGAGGGTCAITTTGGCGGAGCAA 822
 |||||

RESULT 13
 BU025918 769 bp mRNA linear EST 23-AUG-2002
 LOCUS OHG12C07.YG.ab1 OH_EFGHU sunflower RHA280 Helianthus annuus cDNA
 clone OHG12C07, mRNA sequence.

ACCESSION BU025918
 VERSION BU025918.1 GI:22461438
 KEYWORDS EST
 SOURCE Helianthus annuus (common sunflower)

ORGANISM Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroidae;
 Heliantheae; Helianthus.
 1 (bases 1 to 769)

REFERENCE Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
 Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
 P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
 Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://comgenomics.ucdavis.edu/>
 Unpublished

JOURNAL
 COMMENT Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig OH_CA_Contig2686, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: OHG12 row: C column: 07.
 Location/Qualifiers

FEATURES
 source
 1..769
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clone="OHG12C07"
 /lab_host="E.coli"
 /clone_lib="OH_EFGHU sunflower RHA280"
 /notes="Vector: pBRCDNASfiAB; The library was constructed
 from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at <http://cgpdb.ucdavis.edu/>
 TAG_SEQ=Not found"

BASE COUNT 221 a 138 c 194 g 216 t
 ORIGIN

Query Match 32.0%; Score 483.2; DB 13; Length 769;
 Best Local Similarity 77.7%; Pred. No. 2.6e-130;
 Matches 584; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 355 GTGGAGTAGGACCGTTCGATCATGGAGCCATAAATGTTGGCCGTTTGGCCCTCAAG 414
 |||||
 Db 18 GTCCGGGTTGGAACTCTGGGCTCATCCGAAGCCATCATGTTGGCCGACTAGCTTTCAA 77
 |||||

QY 415 CGTAAATGGCAACAAAGCGCAAAAGCTGAAAGGCAACCCGTCGATAAACCACCAATTGTC 474
 |||||

78 AGAAATGGCAAAACAAATGAAGCTTTGGCAAACTTTGGGACAAACCTAACATTGTA 137
 475 ACCGGAGCCAAATGTTCAAGTGTGTTGGGAAATTCCTAGGTACTTTGAGGTGAACTT 534
 138 ACCGGGCCAAACCTTCAGGTTTGTGGGAAATTTGCTGGTATTTTGAAGTGGAGTTG 197
 535 AAGGAGTGAATGATGGAAGACTACTACTGATGAGACCTTCAACAGCTTTGATATG 594
 198 AAGGAATGGAATGAGGAAAGTTACTACTGATGAGTCCCTGAGAAAGCTTTGGAGAT 257
 595 GTTGATGAGAACACCAATTTGTTGGGACATCTTGGTTCCACTCTTAATGAGAAATC 654
 258 GTTGAUGAAACACTATTTGTGGCTGCTATCTTGGTTCCACCTCAATGGTGAATTT 317
 655 GAAGATTTAAACTTTGAACGATCTCTTGGTCCGAAAGAAACAAAGAAACCCGATGGAT 714
 318 GAAGATTTAAGCGTTTAAATGACCTCTTGATTGAGAAATAATGCCAAGCCGGATGGAT 377
 715 ACACCAATCCACGTGGATGGGCAAGTGGAGATTCATTGACCGTTTGTATCCGGAA 774
 378 ACACCTATTGATGGGACGCTGCAAGTGGAGGTTTATTGCCACCAATTTATACCCGAA 437
 775 TTGGAATGGGACTTTAGACTTCCCTGTTGGAAGATCAATGTGATGCTCACAAGTAT 834
 438 CTGAAATGGATTTCCGGTTGCCATTTGTTAAGATATAAATGTTAGTGTCAAAATAT 497
 835 GGAATGTTGACGAGGATTTGTTGGTGGATCTGGGAAACAAAGAGGATTTGCCCTGAG 894
 498 GGGCTTGTATCCGCGGAAATTTGGTGGTCAATTTGGAGAAACAAAGATGACTTGCCTGAT 557
 895 GAATCATCTCCATCAATTTCTTGTGCTGCGACCAACCCACTTACTCAATTTTC 954
 558 GAATCATCTTTCACATCAACTATCTTGTGGTGGTATCAACCACTTTCACCTCAACTT 617
 955 TCCAAAGGTTCAAGTCAAGTCAATTTGCTCAACTTACCACCACTTATCCGATTTGGCCACGAG 1014
 618 TCCAAAGGTTCTAGTCAAAATTTGCTCAACTATCACTGATTTGCTTGGTGGTGGAG 677
 1015 GGTTCAGAAATGATGAGGAAATTCAGAGAAATATGATCTCTCAAGGAGGACTT 1074
 678 GGATCAAGAGCATCATGGGAAACTGTCAGAAATAATGCAAAATGTTCTAAAGAAAGTTTG 737
 1075 GAGAGACAGAAGGTTCAACTGCTCTCAA 1106
 738 GAGAAACCCGAGGTTCAACATCACTCAA 769

BU023643 750 bp mRNA linear EST 23-AUG-2002
 QHF12C03.yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
 clone QHF12C03, mRNA sequence.
 BU023643
 BU023643.1 GI:22459163
 EST.
 Helianthus annuus (common sunflower)
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 1 (bases 1 to 750)
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
 Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
 P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
 Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished
 CONTACT: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Amundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig OH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: QHF12 row: C column: 03.
 Location/Qualifiers
 1..750
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /culivar="RHA280"
 /db_xref="taxon:4232"
 /clone="QHF12C03"
 /lab_host="E.coli"
 /clone_lib="OH_EFGHJ sunflower RHA280"
 /note="Vector: pBRCDNASFIAB; The library was constructed
 from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformants made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_SEQ=Not found"

BASE COUNT 221 a 132 c 187 g 210 t
 ORIGIN
 Query Match 32.0%; Score 482.8; DB 13; Length 750;
 Best Local Similarity 77.7%; Pred. No. 3.4e-130;
 Matches 593; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
 QY 361 FTAGAAACCGTTGGATCATCGGAGCCATAATGTTGGCCGTTTGGCCCTTCAAGCGTAAA 420
 Db 1 GTTGAACCTGTGGCTCATCCGAGCCATCATGTTGGCCGACTAGCTTTCAAAAGAAAA 60
 QY 421 TGGCAGAACACCGCAAGCTGAAGCGCAACCCGTCGATAAACCCCAACATGTCACCGGA 480
 Db 61 TGGCAAAACAAATGAAAGCTCTTGGCAAACTTGGGACAAACCTAACATTTAAACCGGG 120
 QY 481 GCCAATGTTCAAGTGTGTTGGAGAAATTCGCTAGTACTTTGAGGTTGAACTTAAGGAA 540
 Db 121 GCCAACGTTTCAGTGTGTTGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGAA 180
 QY 541 GTGAAATGAGTGAAGGATCATATGTGATGGACCCCTCAACAGCTGTTGATATGTTGAT 600
 Db 181 GTGAAGTTGAGGAAAGTTTACTACGTGATGGATCTGAGAAAAGCTGTGGAGATGTTGAT 240
 QY 601 GAGAACACCAATTTGTTGGGACATCTTGGTTCACCTCTTAAATGGAGAAATTCGAAAGAT 660
 Db 241 GAAAAACATTTTGTGTGGCTGCTATCTTGGGTTCCACCCCTCAATGGTGAATTTGAAGAT 300
 QY 661 GTTAAACTCTTGAACGATCTCTTGGTCCAAAGAAACAAAGAAACCCGATGGGATACCCA 720
 Db 301 GTTAGCGTTTAAATGACCTCTTGAATGAGAAATAATGCCAAGCCGGATGGATACACT 360
 QY 721 ATCCACGTGATCGGCAAGTGGAGGATTCATGTGACCCGTTTGTATCCGGAATTTGAA 780
 Db 361 ATTATGTGACGCTGCAAGTGGAGGTTTATTTGACCACTTATTTACCGGAACTTGA 420
 QY 781 TGGGACTTTAGACTTCCCTTGGTGAAGATATCAATGTGAGTGTTCACAAAGTATGGACTT 840
 Db 421 TGGGATTTCCCGTTCCCAATGGTTAAGAGTATAAATGTTAGTGGTTCACAAAATGGGGCT 480
 QY 841 GTGTACGGAGGATTTGTTGGTGTATCTGGAGAAACAAAGAGGATTTCCCTGAGGAACTC 900
 Db 481 GTTTAGCCCGAAATTTGGTGGTCAITTTGGAGAAACAAAGATGACTTCCCTGATGAACTC 540
 QY 901 ATCTCCATATCAATATCTTGGTGTGCTGACCAACCCACTTTTACTCTCAATTTCTCCAAA 960
 Db 541 ATCTTTCACTCACTTCTTGGTGTCTGATCAACCACTTTCACCTCTCACTTCTCCAAA 600
 QY 961 GGTTCAGTCAAGTCAATGCTCAATACCAACTTATCCGATTTGGCCACGAGGGTTAC 1020

FEATURES
 source

601 GFTTCTAGTCAAAATAATGCTCAATACTATCACTAGTTCATTTCGCTGGGTTCCGAGGGATAC 660
 1021 AGAAATGCTGATGAGAAATTCAGAGAGATATGATCGTCTAAGGGAAGGACTTGAGAAG 1080
 661 AGAGCATCATGGAARAAGTCTCAAGAAATGCAAAATGTTCTAAGAAGAAGTTGGAGAAA 720
 1081 ACAGAAAGGTTCAACAFGCTCAAAAGGC 1110
 721 ACCGACGGGTCAACATCATCTTAANGAC 750

Db 1 GTTGGAACTGGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAGAAAA 60
 QY 421 TGGCAGAAACAAGCGCAAAGCTGAAGGCAAAACCCGTCGATAAACCCCAACCAATTTGTCAACCGGA 480
 Db 61 TGGCAAAAACAATAAAGAACTCTTGGCAAAACCTTGGCAAAACCTTAACATTTGTAACCGGG 120
 QY 481 GCCAATGTTCAAGTGTGTTGGAGAAAATTCGCTAGTACTCTTGGAGTTGAACTTAAGGAA 540
 Db 121 GCCAAGCTTCAAGTGTGTTGGAGAAAATTCGCTAGTACTCTTGGAGTTGAACTTAAGGAA 180
 QY 541 GTGAAATGAGTGAAGGATACTATGTGATGACCCCTCAACAAGCTGTGATATGATGGTTCAT 600
 Db 181 GTGAAGTTGAGGAAAGGTTACTACGTGAAGGATCTGAGAAAAGCTGTGAGATGGTTGAT 240
 QY 601 GAGAACACCAATTTGTTGGGACATCTTGGTTCCTCACTTAATGGAGAATTCGAAGAT 660
 Db 241 GAAACACTATTTGTGTGCTGCTATCTTGGTTCACCCCTCAATGGTGAATTTGAAGAT 300
 QY 661 GTTAAACTCTTGAACGATCTCTTGGTTCGAAAAGAAACAAGAAACCGGATGGGATACACA 720
 Db 301 GTTAAAGCGTTTAAATGACCTCTTGAATGAGAAAATGCAAGACCGGATGGGATACACCT 360
 QY 721 ATCCACCTGGATCGGCAAGTGGAGGATTCATTCGACCGTCTTGTATCCGGAATTTGAA 780
 Db 361 ATTCATGTGGAGCGTCAAGTGGAGGTTTATTTGACCACTTATTTACCCGGAATTTGAA 420
 QY 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGTCAACAAGTATGGACTT 840
 Db 421 TGGGATTTCCGGTTCCCAATGGTAAAGATATAAATGTTAGTGTCAACAATATGGGCTT 480
 QY 841 GTGTACGAGGAAATTTGGTGTGATCTGGAGAAAACAAGAGGATTTCCCTGAGGAATTC 900
 Db 481 GTTTACGGCCGAAATTTGGTGGTCAATTTGGAGAAACAAGAGTACTTCCCTGATGAACTC 540
 QY 901 ATCTTCCATATCAATTAATCTTGGTCTGACCAACCCACTTTTACTCTCAATTTCTCCAAA 960
 Db 541 ATCTTTTCAATCAACTATCTTGGTCTGATCAACCACTTTCACTCTCACTCTCAACTTTCCAAA 600
 QY 961 GGTTCAAAGTCAAGTCAATGCTCAATFACCAACTTATCCGATTTGGCCACGAGGTTAC 1020
 Db 601 GGTTCAGTCAAAATAATGCTCAATFACCTATCACTTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660
 QY 1021 AGAAATGATGGAGAAATTCAGAGAAATGATGATCGTCTTAAGGGAAGGACTTGAGAAG 1080
 Db 661 AAGAGCATCATGGAARAAGTCTCAAGAAAATGCAAAATGTTCTTAAGAAGAAGTTGGAGAAA 720
 QY 1081 ACAGAAAGGTTCAACATCGTCTCAAA 1106
 Db 721 ACCGACGGTTCACATCATCTTAA 746

Search completed: October 22, 2003, 13:34:35
 Job time: 3001 secs

SULT 15
 BUS 025264
 DEFINITION
 BU025264 clone QHF8J20, mRNA sequence.
 BU025264.1 GI:22460784
 EST.
 Helianthus annuus (Common sunflower)
 Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 1 (bases 1 to 746)
 Kozik, A., Micheltore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison
 P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
 Church, S., Jackson, L. and Bradford, K.
 Lettuce and sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig2586, see http://cgdb.ucdavis.edu/
 for details.
 Plate: QHF8 row: J column: 20.
 Location/Qualifiers
 1. 746
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clone="QHF8J20"
 /lab_host="E.coli"
 /clone_lib="QH_EFGHJ sunflower RHA280"
 /notes="Vector: pBRCNAGSfiAB; The library was constructed
 from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at http://cgdb.ucdavis.edu/
 TAG_LIB=QH_EFGHJ sunflower RHA280
 TAG_TISSUE=hulls
 TAG_SEQ=GCTAGTCCGGG"
 219 a 131 c 185 g 211 t
 IGIN
 SE COUNT 31.9%; Score 482; DB 13; Length 746;
 Query Match 77.9%; Pred. No. 5.8e-130;
 Best Local Similarity 0; Mismatches 165; Indels 0; Gaps 0;
 Matches 581; Conservative 0;

361 GTAGAACCGTTGGATCATCGGAGCCATATGTTGGCCCGGTTGGCCITCAAGCGTAAA 420

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M nucleic - nucleic search, using sw model
un on: October 22, 2003, 12:44:32 ; Search time 425 seconds
(without alignments)
9521.971 Million cell updates/sec

US-10-006-852-1
effect score: 1509
equene: 1 atgggtctcccagcgcgt.....agacgagtggtatctgctaa 1509

coring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
sarched: 1792395 seqs, 1340900451 residues
otal number of hits satisfying chosen parameters: 3584790

inimum DB seq length: 0
aximum DB seq length: 2000000000
ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- atabase : Published Applications,NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: sult No., Score, % Match, Query Length, DB ID, Description. Contains 17 rows of sequence data.

Table with columns: sult No., Score, % Match, Query Length, DB ID, Description. Contains 45 rows of sequence data.

ALIGNMENTS

RESULT 1
US-10-006-852-1
; Sequence 1, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
; APPLICANT: Kinnersely, Alan M.
; APPLICANT: Turano, Frank J.
; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
; FILE REFERENCE: 7224-65
; CURRENT APPLICATION NUMBER: US/10/006,852
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/246,367
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1509)
; OTHER INFORMATION:
US-10-006-852-1

Table with columns: Qy, Db, Qy, Db, Qy. Contains alignment scores and sequence identifiers for various sequences.

b 121 GCGGCGTATCAGATCATCAACGACGAGTGTGCTTGCAGGGAATCCAGGTTGAACCTTA 180
 y 181 GCCTCCCTTTGTGACGACATGGATGGAGCCTGTGATGATAAATCTCATATGTCCTCCATC 240
 b 181 GCCTCCCTTTGTGACGACATGGATGGAGCCTGTGATGATAAATCTCATATGTCCTCCATC 240
 y 241 AACAGAACTATGTTGACATGACGACGAGTACCCCGTCAACCACCGAATCTCAGAACCGATGT 300
 b 241 AACAGAACTATGTTGACATGACGACGAGTACCCCGTCAACCACCGAATCTCAGAACCGATGT 300
 y 301 GTGAACATGATTCGACATCATCTATTCAATGACCCGTTAGAGAGCGGAGACCCCGTCCGGA 360
 b 301 GTGAACATGATTCGACATCATCTATTCAATGACCCGTTAGAGAGCGGAGACCCCGTCCGGA 360
 y 361 GTAGGAAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCCTTCNAGGGTAAA 420
 b 421 TGGCAGAACAAAGCGCAAGCTGAAGCCAAACCCCGTCTGATAAACCACCAATTTGTCAACCGGA 480
 y 421 TGGCAGAACAAAGCGCAAGCTGAAGCCAAACCCCGTCTGATAAACCACCAATTTGTCAACCGGA 480
 y 481 GCCAATGTTCAAGTGTGTTGGAGAAATTCGCTTAGTACTTTTGGATTTGAACTTAAGAA 540
 y 481 GCCAATGTTCAAGTGTGTTGGAGAAATTCGCTTAGTACTTTTGGATTTGAACTTAAGAA 540
 y 541 GTGAAATTTGAGTGAAGGATACTATGATGAGGACCTCAACAAAGCTGTTGATGTTGAT 600
 y 541 GTGAAATTTGAGTGAAGGATACTATGATGAGGACCTCAACAAAGCTGTTGATGTTGAT 600
 y 601 GAGAACCCATTTGTTGGGACATCTTGGTTCACCTTTAATGGAGAAATTCGAAGAT 660
 y 601 GAGAACCCATTTGTTGGGACATCTTGGTTCACCTTTAATGGAGAAATTCGAAGAT 660
 y 661 GTTAAACTTTGAAACGATCTTGTGTCGAAAGAAACAAAGAAACCGGATGGGATACCA 720
 y 661 GTTAAACTTTGAAACGATCTTGTGTCGAAAGAAACAAAGAAACCGGATGGGATACCA 720
 y 721 ATCCAGTGTGATCGGCAAGTGGAGATTCATTTGACCCGTTTGTGATCCGGAATTTGAA 780
 y 721 ATCCAGTGTGATCGGCAAGTGGAGATTCATTTGACCCGTTTGTGATCCGGAATTTGAA 780
 y 781 TGGACTTTAGACTTCCCTTGGTGAAGAGATCAATGAGTGGTCAACAAATGAGATTTGACTT 840
 y 781 TGGACTTTAGACTTCCCTTGGTGAAGAGATCAATGAGTGGTCAACAAATGAGATTTGACTT 840
 y 841 GTGTACGAGGGATTTGGTGGTGTCTGAGAAACAAAGAGGATTTGCTGAGGAATTC 900
 y 841 GTGTACGAGGGATTTGGTGGTGTCTGAGAAACAAAGAGGATTTGCTGAGGAATTC 900
 y 901 ATCTTCCATATCAATTTCTGTTGTCGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
 y 901 ATCTTCCATATCAATTTCTGTTGTCGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
 y 961 GGTTCAGTCAAGTTCATTAATGCTCAACTACTCAACTTATCCGATTTGGCCACGAGGGTTAC 1020
 y 961 GGTTCAGTCAAGTTCATTAATGCTCAACTACTCAACTTATCCGATTTGGCCACGAGGGTTAC 1020
 y 1021 AGAAATGTATGGAGAAATTCAGAGAAATGATCGTCCGTAAGGAGGACTTGGAGAG 1080
 y 1021 AGAAATGTATGGAGAAATTCAGAGAAATGATCGTCCGTAAGGAGGACTTGGAGAG 1080
 y 1081 ACAGAAAGTTCAACTGCTCTCAAGGACGAGGGAGTCCACTTGTCTTCTCCCTT 1140
 y 1081 ACAGAAAGTTCAACTGCTCTCAAGGACGAGGGAGTCCACTTGTCTTCTCCCTT 1140
 y 1141 AAAGATAGCAGTGTCACTAGTTCGAAATCTCCGACATGCTTCGAGGATGGATGG 1200
 y 1141 AAAGATAGCAGTGTCACTAGTTCGAAATCTCCGACATGCTTCGAGGATGGATGG 1200
 y 1201 ATAGTGGCGGCTTACAAATGCTCCAAATGCAACACATCACTGTTCTTCTGTTGGTT 1260
 y 1201 ATAGTGGCGGCTTACAAATGCTCCAAATGCAACACATCACTGTTCTTCTGTTGGTT 1260

Qy 1261 ATCAGAAAGATTTCTCGAARACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 Db 1261 ATCAGAAAGATTTCTCGAARACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 Qy 1321 ATGCGTGGACTCGATGAGCTTCTTCGAGAGAGTGTTCACAAAATATCACTTTGGACAAGAG 1380
 Db 1321 ATGCGTGGACTCGATGAGCTTCTTCGAGAGAGTGTTCACAAAATATCACTTTGGACAAGAG 1380
 Qy 1381 AAGAGTGAATCTAAACAGCGATTAACCTTGTGATGATGCTGAGAGAGCGGATATCGACAAAG 1440
 Db 1381 AAGAGTGAATCTAAACAGCGATTAACCTTGTGATGATGCTGAGAGAGCGGATATCGACAAAG 1440
 Qy 1441 CAGAGAGATATCATCACTGGCTGGAGAAAGTGTTCGCGCACAGGAAAGACCGAGTGGT 1500
 Db 1441 CAGAGAGATATCATCACTGGCTGGAGAAAGTGTTCGCGCACAGGAAAGACCGAGTGGT 1500
 Qy 1501 ATCTGCTAA 1509
 Db 1501 ATCTGCTAA 1509

RESULT 2
 US-09-938-842A-1810 ; Sequence 1810, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1810
 ; LENGTH: 1509
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1810

Query Match 99.9%; Score 1507.4; DB 10; Length 1509;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGTGTCTCCACACCCCGTATCGGAGTCGGACGCTCCGTCCTCCACTCCACATTCGCATCA 60
 Db 1 ATGGTGTCTCCACACCCCGTATCGGAGTCGGACGCTCCGTCCTCCACTCCACATTCGCATCA 60
 Qy 61 CGTTAGTCCGTTACTTCACTTCTTAGGTTCAAGATGCGCGAAACTCGATTCCCTAAGGAA 120
 Db 61 CGTTAGTCCGTTACTTCACTTCTTAGGTTCAAGATGCGCGAAACTCGATTCCCTAAGGAA 120
 Qy 121 GCGCGGTATCAGATCATCAACGACGAGCTGATGCTTGAACGGAATCCACGTTGAACCTTA 180
 Db 121 GCGCGGTATCAGATCATCAACGACGAGCTGATGCTTGAACGGAATCCACGTTGAACCTTA 180
 Qy 181 GCCTCTTTGTGACGACATGGATGGAGCCTGTGATGATAAACTCATGTCCTCCATC 240
 Db 181 GCCTCTTTGTGACGACATGGATGGAGCCTGTGATGATAAACTCATGTCCTCCATC 240
 Qy 241 AACAGAACTATGTTGACATGACGAGTACCCCGTCAACCACCGAATTCAGAACCGATGT 300
 Db 241 AACAGAACTATGTTGACATGACGAGTACCCCGTCAACCACCGAATTCAGAACCGATGT 300

301 GTGAACATGATTGCACATCTATTCAATGACCGTTAGAGAGCGGAGACCGCGTCGGA 360
 b GTGAACATGATTGCACATCTATTCAATGACCGTTAGAGAGCGGAGACCGCGTCGGA 360
 361 GTAGAACCCGTTGGATCAICGAGAGCCATAAATGTCGCGGTTGGCCCTCAAGGTTAA 420
 b GTAGAACCCGTTGGATCAICGAGAGCCATAAATGTCGCGGTTGGCCCTCAAGGTTAA 420
 421 TGGCAGAACAGCCAAAGCTGAGGCAACCCGTCGATAAACCNAACHTTGTCACCGGA 480
 b TGGCAGAACAGCCAAAGCTGAGGCAACCCGTCGATAAACCNAACHTTGTCACCGGA 480
 481 GCCAATGTTCAAGTGTGGGAGAAATTCGCTAGGTACTTTGAGTTGAACTTAAGGAA 540
 b GCCAATGTTCAAGTGTGGGAGAAATTCGCTAGGTACTTTGAGTTGAACTTAAGGAA 540
 541 GTGAATTTGAGTGAAGACTGATGAGGACCTCAACAGCTGTGATGTTGAT 600
 b GTGAATTTGAGTGAAGACTGATGAGGACCTCAACAGCTGTGATGTTGAT 600
 601 GAGAACACCAATTTGTTGGGACATTTCTGTTCCACTTTAATGGAGATTCGAAGAT 660
 b GAGAACACCAATTTGTTGGGACATTTCTGTTCCACTTTAATGGAGATTCGAAGAT 660
 661 GTTAAACTTTGACATCTTTGGTTCGAAAGAACAAAGAAACCGGATGGGATACCCA 720
 b GTTAAACTTTGACATCTTTGGTTCGAAAGAACAAAGAAACCGGATGGGATACCCA 720
 721 ATCCAGTGTGAGGAGGAGGATTCATTCACCCGTTTTGTTGATCCGGAAATGGAA 780
 b ATCCAGTGTGAGGAGGAGGATTCATTCACCCGTTTTGTTGATCCGGAAATGGAA 780
 781 TGGACTTTAGACTTCCCTTTGGTGAAGATCAATGTGAGTGTGACAAAGTATGGACT 840
 b TGGACTTTAGACTTCCCTTTGGTGAAGATCAATGTGAGTGTGACAAAGTATGGACT 840
 841 GTGTACCCAGGATTTGGTGGTGAATTCGAGAACAAAGAGGATTTGCTCAGGAACTC 900
 b GTGTACCCAGGATTTGGTGGTGAATTCGAGAACAAAGAGGATTTGCTCAGGAACTC 900
 901 ATCTTCCATATCAATATCTTGTGCTGACCAACCCACTTTACTCTCAATTTCCAAA 960
 b ATCTTCCATATCAATATCTTGTGCTGACCAACCCACTTTACTCTCAATTTCCAAA 960
 961 GGTTCAGTCAAGTCAATGCTAACTACCAACTATCCGATTTGGCCACAGGGTTAC 1020
 b GGTTCAGTCAAGTCAATGCTAACTACCAACTATCCGATTTGGCCACAGGGTTAC 1020
 1021 AGAATGTGATGGAGAAATTCAGAGAAATGATGCTCCTAAGGAGGAACTTGAAG 1080
 b AGAATGTGATGGAGAAATTCAGAGAAATGATGCTCCTAAGGAGGAACTTGAAG 1080
 1081 ACAGAAAGTTCAACATGCTCAAGGAGGAGTGCCACTTGTGCTTCCCTTG 1140
 b ACAGAAAGTTCAACATGCTCAAGGAGGAGTGCCACTTGTGCTTCCCTTG 1140
 1141 AAGATAGACGCTGACACTGAGTTCGAAATCTCCGACATGCTTCCGAGGATCGATGG 1200
 b AAGATAGACGCTGACACTGAGTTCGAAATCTCCGACATGCTTCCGAGGATCGATGG 1200
 1201 ATAGTGGCGGCTACAAATGCGCTCCTCAAAATGCACAAACATCACTGTTTCCGTTG 1260
 b ATAGTGGCGGCTACAAATGCGCTCCTCAAAATGCACAAACATCACTGTTTCCGTTG 1260
 1261 ATCAGAGAAGTTCTCGAGAACACTGCTGAGAGACTTGTGATCGAATATAGAGAAAGTG 1320
 b ATCAGAGAAGTTCTCGAGAACACTGCTGAGAGACTTGTGATCGAATATAGAGAAAGTG 1320
 1321 ATGCGTGAAGTTCGATGAGTTCCTTCCGAGAGTGAATCACAATAATCACTTGGACAAG 1380
 b ATGCGTGAAGTTCGATGAGTTCCTTCCGAGAGTGAATCACAATAATCACTTGGACAAG 1380
 1381 AAGAGTGAATCTAACAGCGAATCTTGTGATGCTCAGGTTGAAGAGCGGATATCGCAAG 1440

Db 1381 AAGAGTGAATCTAACAGCGAATTAATGATGGTTCACCGTGAAGAGCGGATATCGCAAG 1440
 Qy 1441 CAGAGATATCACTACTGCTGGAAAGAAAGTTTTCGCGACAGGAAGAAAGACGAGTGT 1500
 Db 1441 CAGAGATATCACTACTGCTGGAAAGAAAGTTTTCGCGACAGGAAGAAAGACGAGTGT 1500
 Qy 1501 ATCTGCTAA 1509
 Db 1501 ATCTGCTAA 1509
 RESULT 3
 US-10-006-852-11
 ; Sequence 11, Application US/10006852
 ; Publication No. US20030046732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kinnersely, Alan M.
 ; APPLICANT: Turano, Frank J.
 ; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 ; FILE REFERENCE: 7224-65
 ; CURRENT APPLICATION NUMBER: US/10/006,852
 ; PRIOR APPLICATION NUMBER: US 60/246,367
 ; PRIOR FILING DATE: 2000-07-01
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 11
 ; LENGTH: 1705
 ; TYPE: DNA
 ; ORGANISM: Nicotiana tabacum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (71)..(1558)
 ; OTHER INFORMATION:
 US-10-006-852-11
 Query Match 58.5%; Score 883.4; DB 14; Length 1705;
 Best Local Similarity 75.2%; Pred. No. 1.2e-281;
 Matches 1125; Conservative 0; Mismatches 351; Indels 21; Gaps 1;
 Qy 1 ATGTGCTCTCCACAGCGCTATCGGAGTGGAGCTCTCCGCTCCACTCCACATTCGCATCA 60
 Db 71 ATGTTCTGTTCCAGACAGCGTCCGAAAGTGGCTCCATCCACTCCACTTTCGCTTC 130
 Qy 61 GFTTACGTCGATCTTCACTTCTTAGGTTCAAGATCCCGAAACCTCGATTCCTAAGGAA 120
 Db 131 CGATATGTTCCGTACTTCTTCCGAGGTTTAAAGTCCAGAAATTCGATACCAAGGAA 190
 Qy 121 GCGGCTATCAGATCAATCAACGACGAGCTGATGCTTTGACGGGAATCCAGGTTGACTTA 180
 Db 191 GCAGCATATCAATCAATATGATGAGCTTATGTTAGATGGAAATCCAGACTAATTTA 250
 Qy 181 GCCTCTTTTGCAGCATGGATGGAGCCCTGATGATGATAAACTCATATGTCCTCCATC 240
 Db 251 GCATCTTTTGTGACCAATGGATGGAAACCAAGAGTGTAAACAACTGATGATGATTCAT 310
 Qy 241 AACAGAACTATGTTGACATGGACGAGTACCCCGTCCACCCGAACTTCAGAACCGATGT 300
 Db 311 AACAGAAATTTACCTGTGACATGGATGAAATACCCCTGTAAACCACTTCAAGATCGATGT 370
 Qy 301 GTGAACATGATTGCACATCTATTCAATGACCGTTAGAGAGCGGAGACCGCGTCGGA 360
 Db 371 GTGAACATGATTGCACATCTATTCAATGACCGTTAGAGAGCGGAGACTCGAGTTGGA 430
 Qy 361 GTAGAACCCGTTGGATCAICGAGAGCCATAAATGTCGCGGTTGGCCCTCAAGGTTAA 420
 Db 431 GTTAGAACTGTTGGATCTCTGAGGCTTATGATGCTTGGATTTAGCTTCAAGAGAAA 490
 Qy 421 TGGCAGAACAGCCAAAGCTGAGGCAACCCGTCGATAAACCNAACHTTGTCACCGGA 480
 Db 491 TGGCAAAATAAATGAAGAGCCCAAGCGCCCTGTGACAGCCCAATATTTGCTACTGT 550


```

721 ATCCACGTGGATCGGCAAGTGGAGGATTCATTGACCCGTTTTTGTATCCGGAATGGAA 780
792 ATTCTATGAGATGACGAGAGTGGTGGATTTATTGACCCGTTCAITTTACCAGAGCTTGG 851
781 TGGACTTTAGACTCCCTTGGTGAAGAGATCAATGTGAGTGGTTCACAAAGTATGGACTT 840
852 TGGGACTTTAGATGCAATAGTGAAGAGCAATTAATGAAAGTGGTTCACAAATATGGCTT 911
841 GTGTACCGCAGGATTTGGTGGTGTCTGAGAGCAAAAGAGAGATTTGCTGAGGAACTC 900
912 GTCTATGCTGATTTGGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 971
901 ATCTTCCATATCAATTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
972 ATCTTCCATATTAATTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1031
961 GGTTCAGTCAAGTCAATGCTCAATCACTCACTCACTCACTCACTCACTCACTCACTCACT 1020
1032 GGTTCAGTCAAGTCAATGCTCAATCACTCACTCACTCACTCACTCACTCACTCACTCACT 1091
1021 AGAAATGTGATGGAGAAATGGCAGAGAAATGATCGTCTAAGGGAAGACTTGAAGAG 1080
1092 AAGAAATGTGATGGAGAAATGGCAGAGAAATGATCGTCTAAGGGAAGACTTGAAGAG 1151
1081 ACAGAAAGTTCAATGCTCAATCACTCACTCACTCACTCACTCACTCACTCACTCACT 1140
1152 ACAGAAAGTTCAATGCTCAATCACTCACTCACTCACTCACTCACTCACTCACTCACT 1211
1141 AAAGATAGCAGTGTCACTGATGCTGAAATCTCCGACATGCTTCGCAAGTATGGATGG 1200
1212 AAAGACACAGGCAACACAGGATTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271
1201 ATAGTGGGGCTTACAAATGCTCCAAATGCAACACATCACTCTTCTTCTGTTGGTT 1260
1272 ATTGTTCTGCATATCTATGCCACCAACGCAACACATTAACATTTCTCAGAGTTGG 1331
1261 ATCAGAGAAGATTTCTGAGAACACCTGCGAGAGACTTGTGATCGATATAGABAAGTG 1320
1332 ATCAGAGAAGATTTCTCCGATGCTGCGAGAGACTTGTGATCGATATAGABAAGTG 1391
1321 ATGCGTGGCTGATGAGTCTTCCGAGAGTGAATTCACAAATATCACTTTGGAACAAGAG 1380
1392 CTTCAATCAACTTGCACACATCTCCGACGCTCAATGCTAAGCTCGCTGTCGCGAGAG 1451
1381 AAGAGTAACTTAACAGGATACTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1440
1452 CAGCGGCTGGATGTC-----AGCGAGTGCATGAAGAAACAGATAGCCGAAGTG 1502
1441 CAGAGATATCATCTGCTGGAGAAATTTGTCGCGACAGGAGAGAGAGAGT 1497
1503 CAGTTGGAGATGATACTGCAATGAGAGAAATTTGTTGAGAAAGAAAGAGAGAGACT 1559

```

```

SUIT 5
-10-006-852-3
Sequence 3, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
APPLICANT: Kinnersely, Alan M.
APPLICANT: Turano, Frank J.
TITLE OF INVENTION: Methods for Regulating Plant GABA Production
FILE REFERENCE: 7224-65
CURRENT APPLICATION NUMBER: US/10/006,852
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: US 60/246,367
PRIOR FILING DATE: 2000-11-07
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 1665
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (17)..(1498)
; OTHER INFORMATION:
US-10-006-852-3

Query Match
Best Local Similarity 57.7%; Score 870.2; DB 14; Length 1665;
Matches 1110; Conservative 0; Mismatches 348; Indels 21; Gaps 1;

Qy 31 GACGTCCTCCGTCACATTCACATTCGCAATTCGATTCAGTTCGTCGCTACCTTCTCTAGGTTTC 90
Db 44 GATGAATCTCTCGCACCATGTTCCGATATCGCTATGTTCCGACATACCTTCCCAAGTAT 103
Qy 91 AAGATCGCGAAACATCGATTCCTTAAGGAGCGGCTATCAGATCATCAACACGAGGCTG 150
Db 104 GAGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 163
Qy 151 ATGCTTACCGGAAATCCACGGTTGAACTTAGCCTCTTTGTGACGCAATFGATGGAGCCT 210
Db 164 ATGCTTACCGGAAATCCACGGTTGAACTTAGCCTCTTTGTGACGCAATFGATGGAGCCT 223
Qy 211 GAGTGTGATAACTCATGCTCCTCCATCAACAAGAACTATGTTGACATGACAGCAGTAC 270
Db 224 GAGTGTGATAACTCATGCTCCTCCATCAACAAGAACTATGTTGACATGACAGCAGTAC 283
Qy 271 CCGTCCACCGAACTTCAGAACCGGATGTGAACTTAGCCTCTTTGTGACGCAATFGATGG 330
Db 284 CCGTCCACCGAACTTCAGAACCGGATGTGAACTTAGCCTCTTTGTGACGCAATFGATGG 343
Qy 331 CCGTCCACCGAACTTCAGAACCGGATGTGAACTTAGCCTCTTTGTGACGCAATFGATGG 390
Db 344 CCGTCCACCGAACTTCAGAACCGGATGTGAACTTAGCCTCTTTGTGACGCAATFGATGG 403
Qy 391 ATGTTGCGCGTTGGCTTCAAGCGCTTAAATGGCAACAAGCGCAAGCTCAAGGCAAA 450
Db 404 ATGTTGCGCGTTGGCTTCAAGCGCTTAAATGGCAACAAGCGCAAGCTCAAGGCAAA 463
Qy 451 CCGTCCACCGAACTTCAGAACCGGATGTGAACTTAGCCTCTTTGTGACGCAATFGATGG 510
Db 464 CCGTCCACCGAACTTCAGAACCGGATGTGAACTTAGCCTCTTTGTGACGCAATFGATGG 523
Qy 511 GCTAGTACCTTTCAGGTTGAACTTAAAGGAAAGTAAATTCAGTGAAGGATACATGTTGATG 570
Db 524 GCTAGTACCTTTCAGGTTGAACTTAAAGGAAAGTAAATTCAGTGAAGGATACATGTTGATG 583
Qy 571 GACCTCAACAAGCTTTGATATGTTGATGATGATGATGATGATGATGATGATGATGATG 630
Db 584 GATCCAGACAAAGCAGCAGAAATGGTAGACGAGAAACAATCTGTGTCGACCCATATTTG 643
Qy 631 GGTTCACCTTTAATGGAGAAATTCGAAAGTGTAAACTCTTGAACGATCTTTGGTCGAA 690
Db 644 GGTTCACCTTTAATGGAGAAATTCGAAAGTGTAAACTCTTGAACGATCTTTGGTCGAA 703
Qy 691 AAGAAACGAGGAAATTCGAAAGTGTAAACTCTTGAACGATCTTTGGTCGAA 750
Db 704 AAGAAACGAGGAAATTCGAAAGTGTAAACTCTTGAACGATCTTTGGTCGAA 763
Qy 751 ATGTCACCGTTTTTGTATTCGCGAAATGGAATGGACTTTAGACTTCCCTTTGGTGAAGAGT 810
Db 764 ATGTCACCGTTTTTGTATTCGCGAAATGGAATGGACTTTAGACTTCCCTTTGGTGAAGAGT 823
Qy 811 ATCAAATGAGTGGTCAAGATGACTTGTGTACGAGGAGTGGTGGTGGTGGTGGTGGTGG 870
Db 824 ATCAAATGAGTGGTCAAGATGACTTGTGTACGAGGAGTGGTGGTGGTGGTGGTGGTGG 883
Qy 871 AGAAACGAGGAAATTCGAAAGTGTAAACTCTTGAACGATCTTTGGTCGAA 930
Db 884 AGAAACGAGGAAATTCGAAAGTGTAAACTCTTGAACGATCTTTGGTCGAA 943
Qy 931 CAACCCACTTTTCACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCAATGCTCAATACTAC 990
Db 944 CAACCCACTTTTCACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCAATGCTCAATACTAC 1003

```


1291 ATTAGAAGATTTCTCCGCACACTAGCGGAGCGACTGGTAAATAGACATTTGAAAAGTC 1350
 1321 ATGGGTGAGCTCGATGAGCTTCCCTTCAGAGTGTATTCACAAAATATCACCTTGGACAAAG 1380
 1351 CTCACGAGCTPAGACACACTTCGGCGAGGGTCAACCGCTAAGCTAGCCGTGGCCGAGCGC 1410
 1381 AAGAGTGAUCTAACAGCGATACTTGAATGGTACCGGTGAAGAGCGGATATCGACAAG 1440
 1411 AATGG-----CAGCGCGTGCATTAAGRAAACAGATAGAGAAGTG 1449
 1441 CAGAGATATCATCACTGGCTGGAAGAAGTTTGTCCCGCAGCAGRAAGAGCGAGT 1497
 1450 CAGCTAGAGATTACTGCAATGGAGAAATTTGTCGTATAGAAAGAAAGACT 1506

SULT 7
 -10-006-852-13
 Sequence 13 Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
 APPLICANT: Kimmersey, Alan M.
 APPLICANT: Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 13

LENGTH: 1771
 TYPE: DNA
 ORGANISM: Nicotiana tabacum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (67)..(1554)
 OTHER INFORMATION:
 -10-006-852-13

Query Match 56.8%; Score 857.8; DB 14; Length 1774;
 Best Local Similarity 74.1%; Pred No. 3.8e-273;
 Matches 1109; Conservative 0; Mismatches 367; Indels 21; Gaps 1;

1 ATGGTCTTCCACGGCGTATCGGATCGGAGTCCGFPCCACTCCACATTCGCATCA 60
 67 ATGGTCTTCCACAGACAGCGTCCGAAAGTGAAGTCCGCTCCGCTTCACTCCACTTTCGCCTCC 126
 61 CGTTACCTCGTACTTCACTTCTTAGGTTCAAGATGCGGAAACTCGATTCCTTAGGAA 120
 127 CGATATGTTTCAACTTCTCTTCCAGGTTTAAATATCCAGAGAAATTCATACCANAAGGAA 186
 121 GCGCGTATCAGATCAATCAACGAGCTGATGCTTTCACGGGAATCCACGGTTGAAGTTA 180
 187 GCAGCATATCAGATTAATATGATGACTTATGTTAGATGGAATCCAGGCTAATTTA 246
 181 GCCTCTTTGACGCAATGATGGAGCCGTGAGTGTGATAAATCATCATGCTCCATC 240
 247 GCATCTTTCGTTTCAACATGATGGAGCCAGAAATGTAATACGTTAATGATGGATTCATT 306
 241 AACAGAACTATGTTGACATGAGTACCCCGTCAACCCGACTTCCAGACCGGATG 300
 307 AACAGAACTACGTTGACATGAGTAAATCCCTGTAACCACTGAGCTTCAGAAATCGATGT 366
 301 GTGAACATGATTCACATCTATTCATGCAACCGTTTGAAGAGGGGAGACCGCGTCCGGA 360
 367 GTAATATGATAGTCAATTTGTTTAAATGACCACTTGGAGATGGAGAGACTGGAGTTGGA 426
 361 GTAGAAACCGTTGGATCATCGGAGCCATTAATGTTGGCCGGTTTGGCCCTTCAAGCGTAAA 420
 427 GTTGGAACTGTTGGATCCTCTGAAGCTTATTATGCTTCTGGAATTAGGCTTTAAGAGAAA 486

QY 421 TGGCAGAACAGCGCNAAGCTGAAGGCAAAACCCGTCGATAAACCCCAACATTTCTCACCGGA 480
 DB 487 TGGCAAAATAAAATGAAGCCCAAGCAAGCCCTTTGATAAGCCCAATATTTCTCACCGGT 546
 QY 481 GCCAATGTTCAAGTGTGTTGGAGAAATTCGCTAGGTACTTGTGAGGTTCAAATTAAGGA 540
 DB 547 GCTPAAATGTCAGGTTGTTGGAGAAATTTGCAAGTATTTTGAAGTGGAGTTGAAGAA 606
 QY 541 GTAAATGAGTGAAGGATATATGTGATGAGCCCTCAACAAAGCTTTGATATGTTGAT 600
 DB 607 GTAAATGAGTGAAGGATATATGTGATGAGCCCTGAGAAAGCTTGGAAATGTTGAT 666
 QY 601 GAGAACACCATTTGTTGCGGACATTTGTTGTTCCACTTCTTAATGGAGAAATTCGAAGAT 660
 DB 667 GAGAATPACCAATTTGTTGCTCTATCTTAGTTCACCACTCAATGGTGAATTTGAAGAT 726
 QY 661 GTTAAACTCTTGAAACGATCTTTGTTGCAAAAAGAAACAAAGAAACCCGGATGGGATACCCA 720
 DB 727 GTTAGCCGTTTGAATGACCTTTTGAATGAGAAAGAAACAAAGAAACCCGGTGGACACTCCA 786
 QY 721 ATCCAGTGGATCGGCAAGTGGAGGATTCATTGGACCCGTTTTTGTATCCGGAATGGAA 780
 DB 787 ATTCAATGCGATGCAAGCAAGTGGTGAATTTATGGACCATTCCTTTATCCAGAGCTTGA 846
 QY 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCAAAAGTATGACTT 840
 DB 847 TGGGACTTTAGATTCGCATTTGGAGAGAGTATAATGTGAGTGGTCAAAAATATGGCTT 906
 QY 841 GTGTAGCGAGGATTTGGTTGGTGTGATTCGGAGAAAACAAAGAGGATTTGCTCAGGAACTC 900
 DB 907 GTCATGCTGGTATTTGGTTGGCCATTTGGAGAAATGAGAAAGACTTTGCTCTGATGACTT 966
 QY 901 ATCTCCATPACAAATATCTTGGTCTGACCAACCCACTTTACTCTCAATTTCTCCAAA 960
 DB 967 ATTTTCCACATCAATTAACCTTGGTCTGATCAACCTACTTTTCACTCTCAACTTCTAAA 1026
 QY 961 GGTTCAGTCAAGTCAATGCTCAATTAACCAACTTCCGATTCGGAATTCGAGCCACGAGGTTAC 1020
 DB 1027 GGTTCAGCCAAAGTAAATGCTCAATATTAACCAACTTATTCGCTTGGTTTTGAGGTTAC 1086
 QY 1021 AGAATGCTGATGGAGAAATTCAGAGAGAAATATGATGCTCTAAGGAGGAGACTTGGAGAG 1080
 DB 1087 AAGAAATGTTATGAGAAATTTGTCAGAAATGCAAGGATTAAGAGAAAGAAATGAAAAA 1146
 QY 1081 ACAGAAAGTTCAACATCGTCTCAAGGACGAGGAGTCCCACTTTCGCTTCTCTTTG 1140
 DB 1147 AGTGAAGATTCACATTAATTCCTCAAGAAATTTGGATTCCTTATAGTAGCAATTTCTCTT 1206
 QY 1141 AAAGATAGCAGCTGTCACTGAGTTCGAAATCTCCGATGCTTCCGATGCTTCCGAGGATGGATGG 1200
 DB 1207 AAAGAACACAGTCAACACAAATGAGTTCGAAATTTCTGAAACTTCTTAGAAGATTTGGATGG 1266
 QY 1201 ATAGTCCCGCTTACACATGCTCCAAATGCAACACACATCACTGTTCTTCTGTTGTT 1260
 DB 1267 ATTTGTTGCTGATATATGATGCAACCAATGCTCAACATGCTCAGAGTCTCAGAGTGTGTC 1326
 QY 1261 ATCAGAGAAAGTTTCGAGAACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 DB 1327 ATTAGAGAAAGTTTCTCCCGCACACTAGCGGAGGACTGTTAATAGACATTTGAAAAGTTC 1386
 QY 1321 ATGCGTGAAGTTCGATGAGGTTTCCCTTGGAGAGTTCACAAAATATCACTTGGCAAGAG 1380
 DB 1387 TTCCACGAGAGTAGACACACTTCCGCGAGGGTCAACGGTAAAGTACCGTGGCCGAGCGC 1446
 QY 1381 AAGAGTGAATCTAACAGCGATAAATTTGATGTCAGGTCGAGAGGAGGATTCGACAAG 1440
 DB 1447 AATGG-----CAGCGCGTGCATAGAAAACAGATAGAGAGTG 1485
 QY 1441 CAGAGAGATCATCACTGGCTGGAGAAAGTTTTCGCCCGCAGGAAAGACGAGT 1497
 DB 1486 CAGCTAGAGATTACTGATGTTGAAATTTTGTGTTGATTAAGAAAGAAAGACT 1542

SOFTWARE: FastSeq for Windows Version 4.10
 SEQ ID NO 782
 LENGTH: 1455
 TYPE: DNA
 ORGANISM: Oryza sativa
 -09-887-576-782

Query Match 47.0%; Score 709.6; DB 10; Length 1455;
 Best Local Similarity 71.8%; Pred. No. 5.1e-224;
 Matches 928; Conservative 0; Mismatches 364; Indels 0; Gaps 0;

/	45	CTCCACATTCGCATACAGTTACGTCGCTACTTCACTTCTAGTTCAGATCCCGGAAA	104
>	42	CTCGAGTTCGCTCGGGTACGTGCGCAGCGGCTCGGAGTTCCAGATCCCGGAAA	101
/	105	CTCGATTCCTAAGAGCGGCGTATCAGATCATCAACGACGACGTGATGTCGGGAA	164
o	102	GTCCATCCCAAGGACCGCGCTACACAGATCATCAACGACGACGTGATGTCGACG	161
y	165	TCCACGGTTGAACTTAGCCCTCTTGTGACGACATGGATGGAGCCTGATGATAACT	224
b	162	CCCGGCGTGAACCTGGCGTCTTCTGTCACACAGTGGATGGAGCCGAGTCCGAC	221
y	225	CATCATGCTCTCCATCMAAAGAACTATGTTGACATGGACGAGTACCCCTCACC	284
b	222	CATGATGGCCGCATCAACAAAGAACTACGTCGACATGGATGATGATACCCCT	281
y	285	ACTTCAGAACCGATGTGTAACATGATTCACATCTATTCAATGACCGTTAGAGAG	344
b	282	GCTCCAGAACCGGTGGTGAACATGATCCGGCATCTGTTCAACGCGCGATCCG	341
y	345	GGAGCCCGCTCGAGTAGGAACCGTGTGATCATCGGAGGCCATAATGTTGGCG	404
jb	342	CGAGACCGCGGTGGGGTGGCCAGTGGGGTGGTGGAGGCCATCATGCTGGCG	401
xy	405	GGCCTTCAAGCTAAATGGCAGAACAAAGCAGAAAGTGAAGCAACCCCTCGAT	464
jb	402	GGCTTCAAGAGNAGTGGCAGAACAGGATGAAGCCGAGGGAGCCCCACGAC	461
Qy	465	CAACATGTCACCGGAGCCAACTTCAAGTGTGTTGGGAGAAATTCGCTAGT	524
Db	462	CAACATCGTGACCGGGCCCAACGTCGAGGTGTGCTGGGAGAAATTCGCGC	521
Qy	525	GGTTGAACTTAAGGAGTGAATTTGATGAAAGATATGATGATGAGCCCTCA	584
Db	522	GGTGGAGCTCAAGGAGTGAAGCTGACCCAAAGGTTACTACGTGATGAAC	581
Qy	585	TGTTGATATGGTTGATGAGAACCACTTGTGTTGGGACATTCCTGTTCA	644
Db	582	CGTGGAGATGGTCGACGAGAACCACTGCGGTGCTGGGAGAAATTCGCGC	641
Qy	645	TGGAGAAATTCGAGATGTTAAACTCTTTGAACGATCTTTTGGTCGAAAGAA	704
Db	642	CGGCGAGTTCCGAGAGCTCAAGATGCTCAACGACCTCTCACCGCCAGAG	701
Qy	705	CGGATGGGATACCAATCCACGTGATCGGAGTGGGAGTGGAGGATTCAT	764
Db	702	AGGTTGGAACACGCGCTCATGGAACGCGGCGAGCCGCGGTTTCATCGC	761
Qy	765	GTATCCGGAAATGGAAATGGACTTTAGACTTCCCTTGGTGAAGAGTAT	824
Db	762	CTACCCGGAGCTGGAGTGGGACTTCCGGCTCGCTGCTGGTGAAGAGCA	821
Qy	825	TCACAAGTATGGACTTGTACCGAGGGATGGTGGGTTGGTGTGAGAGAA	884
Db	822	CCACAAGTATCGGGTCTCTACGCGCGCTCGGGTGGTCACTGGCGCAAC	881
Qy	885	TTTGCTGAGGAACTCATCTCCATATCAATATCTTGGTGTGACCAACCC	944
Db	882	CCTCCCGATGAGCTCATCTCCACATCAACTACCTCGGCGCGACCCAC	941
Qy	945	TCTCAATTTCTCCAAAGGTTCAAGTCAAGTCAATGTTCAATACCAACT	1004

Db	942	GCTCAACTTCTCCAAAGGATCGAACAGATAATTGGCGAGTATFACCACCT	1001
Qy	1005	GGGCCACGAGGGTTACAGAAATGTGATGGAGAAATTCAGAGAGAAATGAT	1064
Db	1002	CGGATTCGAGGGGTACAAGGACATCATGCAAGAACTCCCGGGACAACCG	1061
Qy	1065	GGAGGACTTGAGAAAGACAGAAAGGTTCAACATCCCTCAAAGGACGAG	1124
Db	1062	GGAGGGATCGAGAAAGACGGGCCACTTTCGTCGTTGGTTCMAGGACT	1121
Qy	1125	TGTCGCTTTCTCCTTGAAGATAGCAGCTGTCACTAGTTCGAAATCTCCG	1184
Db	1122	GGTGGCTTCTCCCTCAAGGACTCGTCGCGTACAGGTTCGAGTGGCGG	1181
Qy	1185	TCCGAGTATGATGATAGTGGGCTACCAATGCTCCAAATGCAACACAT	1244
Db	1182	CCGCGCTTCCGGCTGGATCGTCCCGGTTACCCCTGCGCGGCTCGCCG	1241
Qy	1245	TGTTCTTCTGTGTGTTATCAGAGAAATTTCTCGAACAACCTCGCTGAG	1304
Db	1242	CGTGTGCGCTCTGTCATCCGCGAGGACTTCAGCCGGGCTCGCCGAG	1301
Qy	1305	CGATATAGAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1336
Db	1302	CGACTCACRAGACGGTGGCCGATATGGACG	1333

Search completed: October 22, 2003, 15:04:15
 Job time : 439 secs

PRIOR APPLICATION DATA: GB 9305860.0
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA: GB 9305862.6
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA: GB 9314351.9
 FILING DATE: 12-JUL-1993
 PRIOR APPLICATION DATA: GB 9320988.0
 FILING DATE: 12-OCT-1993
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1783 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE: ERTD1
 ORGANISM: ERTD1
 US-08-522-421-8

Query Match 48.8%; Score 736.6; DB 2; Length 1783;
 Best Local Similarity 71.8%; Pred. No. 4.2e-232;
 Matches 964; Conservative 0; Mismatches 379; Indels 0; Gaps 0;

QY 20 TATCGAGTGGAGCGTCCGCTCCACTCCACATTCGCATCGCTACGTTACGTCGGTACTTCCAC 79
 DB 28 TAAGAGATTGAGAGAGAGCTTGCATCTGATTTGCAATGATGATGATGATGATGATGATGAT 87
 QY 80 TTCTTGTTCAGATGCGGMAAATCGAATTCCTAAGGAAGCGGCGTATCGATCATCA 139
 DB 88 TACTTAAGTTCAAATGCTCTAAATAATCCATCCGAAAGAGCAGCTTATCAGATTTGAA 147
 QY 140 ACGACGAGCTGATGCTGTACGGGAATCCAGCTTGAATCTTAGCCCTCCCTTTGTGCGGACAT 199
 DB 148 ACGACGAGCTGATGCTGTACGGGAATCCAGCTTGAATCTTAGCCCTCCCTTTGTGCGGACAT 207
 QY 200 GATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
 DB 208 GATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
 QY 260 TGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 319
 DB 268 TGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
 QY 320 TATTCAATGACCGTTAGAGAGCGGAGACCGCCGCTCGGAGTAGGAACCCGTTGGATCAT 379
 DB 328 TTTTCCATGCCCCGGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
 QY 380 CGGAGCCATAATGTTGGCCGTTTGGCCCTTCAGGCTAAGGTAAGTGGCAGACAGCCGAAAG 439
 DB 388 CAGAGGCAATAATGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
 QY 440 CTGAAGCAAAACCCGCTGATAAACCACCAATGCTACCGGAGCCAAATGTTCAAGTGTGT 499
 DB 448 CAGAGGCAAAACCCGCTGATAAACCACCAATGCTACCGGAGCCAAATGTTCAAGTGTGT 507
 QY 500 GGGAGAAATTCGCTAGGTAATTTGAGGTTGAACTTAAGGAAGTGAATTTGAGTGAAGGAT 559
 DB 508 GGGAGAAATTTGCAAGGTAATTTGAGGTTGAACTTAAGGAAGTGAATTTGAGTGAAGGAT 567
 QY 560 ACTATGATGAGGACCCCTCAACAGCTGTTGATATGTTGATGAGACACCATTTGTTGTTG 619
 DB 568 ACTATGATGAGGACCCCTCAACAGCTGTTGATATGTTGATGAGACACCATTTGTTGTTG 627
 QY 620 CGGACATTTGTTGCTCCACTCTTAATGAGAAATTCGAAGATGTTAACTCTTGAACCGATC 679
 DB 628 CTGCAATCTTGGTTCTACTCTGACTGGGAGTTTGGAGTGTGAGGATGTTGAGGATGTTGAGG 687
 QY 680 TCTTGGTCGAAAGAAACAAGAACCGGATGGGATGATGATGATGATGATGATGATGATGATGATGAT 739

DB 688 TCCTTACAAAAAAGAAACAAGAAACCCGATGGAGACACCCGATTCATGTGATGCTCGCA 747
 QY 740 GTGAGGATTCATTCACCGTTTTTGTATCGGAATTCGAAATGGACTTTAGACTTCCT 799
 DB 748 GTGGAGATTTATGCT 807
 QY 800 TGTGAAGAGTATCAATGTGAGTGTGTCAAGATGTGGACTTGTGTGACGAGGATGGTT 859
 DB 808 TTTGTAAAAAGTATAAATGTGACGGGTCACAAGTATGGCTTGTATATGCTGCTGCTCGGTT 867
 QY 860 GGGTATCTGGAGAAACAAGAGGATTTCCCTGAGAACTCATCTTCCATATCAATATC 919
 DB 868 GGGTATATGGCGAGCAAGAAAGACTTGGCCGATGAACTGCTTTCATATAAATAC 927
 QY 920 TTGGTCTGACCAACCCACTTACTCTCAATTTCTCCAAAAGGTTCAAGTCAAGTCAATG 979
 DB 928 TTGGGTCIGATCAGCCTACTTTTACTCTCAACTTCTCTAAAGGTTCCCTATCAATAATTG 987
 QY 980 CTCATACTACCACTTATCCGATTTGGCCACGAGGTTACAGAAATGTGATGAGAAAT 1039
 DB 988 CACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGTTATAAAGAACCTCANGAAGAA 1047
 QY 1040 GCAGAGAGAAATATGATCGTCCCTAAGGGAAGACTTGAAGAAGCAAGAAAGTTCAACATCG 1099
 DB 1048 GCTTATCAAAACGCAAAAGTACTTAAACAGAGGATACAAATAATGGGGCGTTTCGATATTG 1107
 QY 1100 TCTCAAAGACGAGGAGTGCCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1159
 DB 1108 TCTCTAAGGATGTTGGGTTCT 1167
 QY 1160 CTGAGTTCCAAATCTCCGACATGTTCCGAGGTATGGATGGATAGTCCGGCTACACAA 1219
 DB 1168 CGTATTGAGGATCTGAGATCTCAGAAATTTGAGTGGATGCTCCCTGCATACACAA 1227
 QY 1220 TGCTCCAAATGCAACACATCACTGTTCTTCTGTTTCTGATATCAGAAAGATTTCTCGA 1279
 DB 1228 TGCCACCGGATGCTGAACACATTTGCTACTCGGGTTGCTATTAGAGAGGATTTTCAGCC 1287
 QY 1280 GAACACTCCCTGAGACTTGTGATCGATATAGAAAGTGTGATGCTGAGCTCGATGAGC 1339
 DB 1288 ACAGCTTAGCTGAGACTTGTGTTCTGCAATGAGAAAATTTCTGTCAGAGTTGGACAC 1347
 QY 1340 TTTCTTCGAGAGTCAATTCACAAA 1362
 DB 1348 AGCTCTCTGTTGCCACCAA 1370

RESULT 2
 US-09-103-840A-2/c
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; TITLE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 4403765
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; FEATURE:
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence
 ; OTHER INFORMATION: represent a, t, c or g
 ; US-09-103-840A-2

Query Match 20.6%; Score 311.2; DB 3; Length 4403765;
 Best Local Similarity 54.1%; Pred. No. 1e-89;
 Matches 687; Conservative 0; Mismatches 568; Indels 15; Gaps 2;

Y 80 TTCCTAGGTTCAAGATCCCGGAAATCTCGATTCCTAAGGAAGCGGGTATCAGATCATCA 139
 b 3837572 TCCGGGCTCGGATCCCGATGAGTCGATGATCCCGAGGGGCTATCGCTTCATCC 3837513

Y 140 ACGAGGAGTGTACTTGAACGGGAATCCAGGTTGAACCTAGCTTCTTGTGACGACAT 199
 b 3837512 ACGACGAGTGTACTTGAACGGGAATCCAGGTTGAACCTAGCTTCTTGTGACGACAT 3837453

Y 200 GGATGGACCTGAGTGTGATAAACTCATATGCTTCCATCAACAAGAACTATGTTGACA 259
 b 3837452 GGATGGACCTGAGTGTGATAAACTCATATGCTTCCATCAACAAGAACTATGTTGACA 3837393

Y 250 TGGAGGATACCCCGTACACACCGAATCCAGAACCGATGTGAAATGATGACATC 319
 b 3837392 AGGACGAATACCCCGGACCGGCGCCATCGAGGGGCTGTGTGTCATGTTGCGCCGAC 3837333

Y 320 TATTCAATGC-----ACCGTTAGAAGCGGGAGACCGCGCTCGGAGTAGGAACCG 370
 b 3837332 TGTTCAGCGCGAGGTTGCGGACACACGCCCCACCGGCGCCCGGGTGTCCACCA 3837273

Y 371 TTGATCATCGGAGCCATATGTTGCGCGGTTTGGCTTCAAGCGTAAATGCGGAAACA 430
 b 3837272 TCGGCTCAGCGAGGCGGTGATGCTGGTGGGCTGCGCTGAAATGGCGTTGGCGGCAAC 3837213

Y 431 AGCGRAAGCTGAAGGCAAAACCCGTCGATAAACCCAAACATGTCACCGGAGCAATGTT 490
 b 3837212 GGTGGGGTCTTGAAGGGGCG-----ATGCCAATCTGGTGTGATGGGTTGCAACGTC 3837159

Y 491 AAGTGTGTGGGAAATTCGCTAGTACTTTGAGGTTGAACCTTAAGGAAGTGAATGA 550
 b 3837158 AGTGTGTGGGAAATTCGCTAGTACTTTGAGGTTGAACCTTAAGGAAGTGAATGA 3837099

Y 551 GTGAAGATACTATGATGATGACCCCTCAACAAGCTGTGATATGTTGATGAGAACACCA 610
 b 3837098 ACCGGGCGGCTACGTATACCCCGGAGAGGTGCTCGCCGCGTFCAGAGAACACCA 3837039

Y 611 TTTGTTGCGGAATCTTGGTTCCTCTTAATGGAGAAATCGAAGATGTTAAACTCT 670
 b 3837038 TCGGCGTGGTGGCAATCTGGGCAACCACTATACCGGTGAACCTCGAAACCCATCGCGAGA 3836979

Y 671 TGAAGATCTCTTGGTCAAAAGAACAAAGAACCCGATGGATACCAATCCAGTGG 730
 b 3836978 TCTGCGCGGCTGGAACAACTGCGGCTGCGGGGTTGGAGCTCCCGTACAGTCC 3836919

Y 731 ATGCGCAAGTGGAGGATTCATGTCACCGTTTTTGTATCCGGAATGGAATGGACTTTA 790
 b 3836918 ACGCGCCAGTGGGGGTTTTGTGTCGCGTTTTTGTATCCGGAATGGAATGGACTTTA 3836859

Y 791 GACTTCCCTGGTGAAGATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 850
 b 3836858 GCCTGCCCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3836799

Y 851 GGAATGTTGGTGTATCGGAGAAACAAAGAGGATTTGCTCGAGGAACTCATCTTCCATA 910
 b 3836798 CGCTCGGTTTTGTCGTTGGCGCGGCGCCGAGACCTCGCGAGGATCTGGTTTTCCGGG 3836739

Y 911 TCATATTCCTGGTGCACCAACCCACTTTACTCTCAATTTCTCAATTTCTCAATTTCT 970
 b 3836738 TCATATTCCTGGTGCACCAACCCACTTTACTCTCAATTTCTCAATTTCTCAATTTCT 3836679

Y 971 AAGTCAITGCTCAATCACTTACCAACTTATCCGATTTGGCCACGAGGTTACAGAAATGCA 1030
 b 3836678 AGTGTGGTGGCCAGTACTACAACTTCTGCGGCTGGGGCGGACGCTATACCAAGTGA 3836619

Y 1031 TGGAAATTCAGAGAAATATGATGCTTCAAGGAGGACTTGGAGAGACGAAGAATG 1090
 b 3836618 TGCAGGCGCTGTGCACACCCCGGTTGGTGGTGTGACCAAGCTGCGGAGGTTGATTCAT 3836559

Y 1091 TCAACATGCTCAAAAGGACGAGGAGTGCACCTTGTGCGTTTTCTCCTTGAAGAATGACA 1150

Db 3836558 GCAGGATGATTCGGATGTTCCGGATCCCGGTTGTCAGCTTCGGCTCCCGCGGAC 3836499

Qy 1151 GCTGTACACTGAGTTTGGAAATCTCCGACATGCTTCCAGATGATGGAATGATGCGCG 1210

Db 3836498 CGGGTACACGGAGTTCGACGCTCCCATGAGTCCGACCTCCGGTGGCAGGTGCCG 3836439

Qy 1211 CCTACCAATGCTCCAAATGCAACACATCACTGTTCTTCTGTTATCAGAGAAG 1270

Db 3836438 CCTACCAATGCTCCAAATGCAACACATCACTGTTCTTCTGTTATCAGAGAAG 3836379

Qy 1271 ATTTCTGAGAAACTCGCTGAGAGACTTGTGATCGATATAGAGAAATGATGCTGAGC 1330

Db 3836378 GACTTCCCGACCTTGGCGCGGCGCTGACGACGACGCGGTCAACCGCTTGGCTGCC 3836319

Qy 1331 TCGATGACT 1340

Db 3836318 TGGACAAGGT 3836309

RESULT 3
 US-09-103-840A-1/c
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE OF INVENTION: TUBERCULOSIS
 ; TITLE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 US-09-103-840A-1

Query Match 20.6%; Score 311.2; DB 3; Length 4411529;
 Best Local Similarity 54.1%; Pred. No. 1e-89;
 Matches 687; Conservative 0; Mismatches 568; Indels 15; Gaps 2;

Qy 80 TTCTTAGGTTCAAGATCCCGGAAATCTCGATTCCTAAGGAAGCGGGTATCAGATCATCA 139
 Db 3851674 TGCAGCGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3851615

Qy 140 ACGACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 199
 Db 3851614 ACGACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3851555

Qy 200 GGAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
 Db 3851554 GGAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3851495

Qy 260 TGGACGAGTACCCCGTACACACCGAATTCAGAACCGATGTTGAAACATGATGACATC 319
 Db 3851494 AGACGAATACCCCGGACCGGCGCTCGAGCGGCTGTGTGTCATGTTGCGCCGAC 3851435

Qy 320 TATTCAATGC-----ACCGTTAGAAGCGGGAGACCGCGCTCGGAGTAGGAACCG 370
 Db 3851434 TGTTCAGCGCGAGGTTGCGGACACACGCCCCACCGGCGCCACCGGGTGTCCACCA 3851375

Qy 371 TTGATCATCGGAGCCATATGTTGCGCGGTTTGGCTTCAAGCGTAAATGCGGAAACA 430
 Db 3851374 TCGGCTCAGCGAGGCGGTGATGCTGGTGGGCTGCGCTGAAATGGCGTTGGCGGCAAC 3851315

Qy 431 AGCGRAAGCTGAAGGCAAAACCCGTCGATAAACCCAAACATGTCACCGGAGCAATGTT 490

```

3851314 GGTGGGGCTTCGGAAAGGGCC-----ATGCCCAATCTGGTGAATGGTTGCAAGCTCC 3851261
491 AAGTGTGTTGGGAAATTCGGTAGTACTTTGAGTTGAACTTTAAGAAAGTGAATTTGA 550
3851260 AGTGTGTGGGAAAGTTCTCCGCTACTTCGAGTGAACCCCGTTACTCTGCCGATGG 3851201
551 GTGAAGGACTATGTGATGGACCTTCAACAACCTGTTGATATGTTGATGAGAACCA 610
3851200 AGCGGGCCGCTACGFCATCACCCCGGACAGGTTCTCCCGCGCTGACGAGAACCA 3851141
611 TTTGTTGTTGGGACATCTTTGTTCCACTTTAATGAGAAATTCGAAGATGTTAAACTCT 670
3851140 TCGCGTGTGGGATCTTTGGCCACCACCTATACGGTGAACCTGAAACCCATCCCGAGA 3851081
671 TGAACGATCTCTGTTGCGAAAGAACAAAGAACCGGATGGATACACCAATCCAGTGG 730
3851080 TCTGCCCGCGCTGGCAAACTGGCGGGTGGCGGGGTGGAGCTCCCGGTACACGTG 3851021
731 ATCCGCAAGTGGAGATTCATTGCACCGTTTTTATCCGGAATTTGAAATGGGACTTTA 790
3851020 ACGCGCCAGTGGGGCTTTGTGTGCCGTTTTTGCATCCGACTGTAATGGGATTTT 3850961
791 GACTTCCCTTGTGAAAGATTCATGTCAGTGTGACAGTGTGACAGTGTGACGAG 850
3850960 GGTGCCCCCGTGTGATCGATCAACGTCAGCGGCCACAAGTATGGGTGACCTATCCC 3850901
851 GGATTTGGTGGTGTCTGGGAAACAAGAGATTTGCCCTGAGGAACTCACTCTCCATA 910
3850900 GCGTGGGTTTGTGTTGCGCGGGCCGAGCACCCTCCGAGGATCTGGTTTTCGGG 3850841
911 TCAATTTCTGTTGTCGACCAACCACCTTTACTCTCAATTTCTCCAAAGTTTCAAGTC 970
3850840 TCAACTACCTCCGCGGGACATGCGCGACCTTCCACCTGAACTTCTCCCGTCCCGTAA 3850781
971 AAGTCAATGCTCAATACTACCAACTTATCCGATTTGGCCAGGAGTTTACAGAAATGTA 1030
3850780 AGTGTGGCCAGTACTACAATCTCTGGGCTGGGCGGAGCTATACCAAGGTGA 3850721
1031 TGGAGAAATGACAGAGATATGATGTCCTTAAAGGAGGACTTGGAGAACAGAAAGT 1090
3850720 TGCAGGCGCTGTGCGACACCGCCCGGTGGCTGGGTGACCACTGCGGAGTGGATCA 3850661
1091 TCAACATGCTCTCAAAGGACGAGGAGTGCACCTTGTGCTTTCTCTGAAAGATAGCA 1150
3850660 GCGAGTGTACTCGGATGTTGCGGATCCCGGTGTCAGCTTCCGGCTCCCGCGGACC 3850601
1151 GCTGTCACTGATGTTGAAATCTCGACATGCTTCCGAGGATGATGATGATGCGCG 1210
3850600 GCGGGTACCGGAGTTCGACGCTTCCCATGACCTCCGAGCTCCGAGCTCCGAGTCC 3850541
1211 CCTACCAATGCTCCAAATGACAAACATCACTGTTCTTCTGTTGTTATCAGAGAAG 1270
3850540 CCTACCAATGCTCCGACAAACCGCAGGAGTGGCGTCTCGGATCGTGGTTCGGAG 3850481
1271 ATTTCGAGAACCTCGCTGAGAGACTTGTGATGATATAGAAAGTGTGCTGAGC 1330
3850480 GACTCTCCCGGACCTGGCGCGGCGCTTGCACGACGACCGCGTCCCGCTGGCTGCC 3850421
1331 TCGATGAGCT 1340
3850420 TGGACAAGGT 3850411

```

```

RESULT 4
US-09-068-195-21
; Sequence 21, Application US/09068195B
; Patent No. 6140078
; GENERAL INFORMATION:
; APPLICANT: Sanders, Jan W.
; APPLICANT: Ledebner, Adrianus M.
; APPLICANT: Venema, Gerard
; APPLICANT: Kok, Jan
; TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid

```

```

; TITLE OF INVENTION: Bacterium, and Its Use in a Lactic Acid Bacterium for
; FILE REFERENCE: Production of Desired Protein
; CURRENT APPLICATION NUMBER: US/09/068,195B
; EARLIER FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: PCT/EP97/04755
; EARLIER FILING DATE: 1997-08-20
; EARLIER APPLICATION NUMBER: EP 97200744/7
; EARLIER FILING DATE: 1997-03-13
; EARLIER APPLICATION NUMBER: EP 96202444/4
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 5565
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genomic DNA of
; OTHER INFORMATION: NS3 locus containing gadRCB
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (3)..(500)
; OTHER INFORMATION: C-terminus of rnhB
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1095)..(1922)
; OTHER INFORMATION: :ggl = gadR
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (2069)..(3577)
; OTHER INFORMATION: orfX = gadC
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (3600)..(4997)
; OTHER INFORMATION: gadB
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: Complement((5078)..(5563))
; OTHER INFORMATION: C-terminus of ORF in opposite direction
; US-09-068-195-21
Query Match 17.9%; Score 270.6; DB 3; Length 5565;
Best Local Similarity 52.4%; Pred. No. 3e-78;
Matches 678; Conservative 0; Mismatches 599; Indels 18; Gaps 3;
Qy 80 TTCCTAGTTCAGATCCCGAAACTCGATCCTCCTAAGGAGCGGCTATCAGATCA 139
Db 3679 TACCTAATATAAATTAGCTCAACAATCAATTGAGCCTCGAGTGGCCTATCAGTTAGTTC 3738
Qy 140 ACGACGAGCTGATCCTTGACGGAAATCCCGTTGAACTTAGCCTCTTTGTGCGACAT 199
Db 3739 AAGATGAAATGCTAGATGAAGGAAACCGCTCGTTAAATTTGGCCACATCTCTCAACTT 3798
Qy 200 GGATGGAGCCTGAGTGTGATAAATCACTCATATGTCCTCCATCAACAGAACTATGTTGACA 259
Db 3799 AATGGAACTTGAAGAGTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3858
Qy 260 TGGACGAGTACCCCGTCCACCACCGAACTTCAGAAACCGATGCTGTGAACATGATTCACATC 319
Db 3859 AATCAGATATCCAAAGAACTGAAATTTGAAACCGTTGCGTCAACATGATCGCTGACC 3918
Qy 320 TATTCATGACCGTTGAGAGCGGAGACCCCGCTCGAGTAGGAACCGTTGATGATCAT 379
Db 3919 TTTTGGAAATGCGAGTGAATAA-----GGAAAAATTTATGGGACTTCGACAAATTTGTTCTT 3972
Qy 380 CGGAGCCATAATGTTGCGCGGTTTGGCTTCAAGCGTAAATGGGAGAACCAAGCGCAAG 439
Db 3973 CAGAGCTTGTATGCTTGGGGATGGGATGATGATGATGATGATGATGATGATGATGATGATG 4032
Qy 440 CTGAAGCAACCCCGTGCAT-----AAACCCAAATTTGTCACCGGAGCCAATGTTCAAG 493
Db 4033 AATTAGGCTAGATATTAATGCGAAAAAGCAAACTTAGTCAATTTTCTGTTTATCAAG 4092

```



```

494 TGTGTTGGAGAAATTCGCTAGGTAATTTGAGTTGAACTTAAGGAAGTGAATTTGAGTG 553
4993 ITTGTCTGGGAAAATTCGTGTGTTTATTTGGGATATTTGAAATGAGAGAGTCCCAATGGATA 4152
554 AAGGATACTATGATGAGACCTCAACAGCTGTTGATATGTTGATGAGAAACCAATTT 613
4153 GAGACATATGTCATCAATTTGGNAAAGTGGATATGTTGATGAAATATACGATTC 4212
614 GTGTTGGGACATCTTGTGTTCCACTTTAATGAGAAATTCGAAGATGTTAAACTTTGA 673
4213 GAGTAGTTGGAATTTGGGGATTAATCTTACTGCTGTTATGATGATCAAAAGCTTTGG 4272
674 ACCATCTTGTGCGAAAAGAAACAAAGGACCGGATGGATACACCAATCCAGTGGATG 733
4273 ATAAATTTGATTTGAAGAAATTAATAACACAGACAGACTACAAGTTTATTTACGATG 4332
734 CGCAAGTGGAGGATTCATTTGCAACCGTTTTGTTATCCGGAAATGGAATGGGACTTTAGAC 793
4333 CTGCTTTCAGGAGACTTTATGCTCTTTTGTGAGCCAGAACTTGGTGGGATTTCCGTT 4392
794 TTCCCTTGTGAGAGATCAATGAGTGTGTCACAAGTATGACTTGTGTACGCGAGGA 853
4393 TGAAAAATGTCATTTCAATCAATCTTCAGACATAAATATGTTTGTAGTATTCCTGGTG 4452
854 TTGTTGGTGTGATCTGGAGAAACAAAGAGGATTTGCTGAGAACTCATCTTCCATCA 913
4453 TAGTTGGTCTTTGGCTGCACAAAAAATATTTACCTGAAAGATTAATTTTTAAAGTAA 4512
914 ATATCTTGTGCTGACCAACCCACTTACTTCAATTTCTCCAAAGTTTCAAGTCAAG 973
4513 GTTAUCTTGAGGAAATTAACCAATGCGGATTAATTTTTCTCACAGTCTTCCAT 4572
974 TCATGCTCAATACCACTTATCCGATTTGCGGACAGGTTTACAGAAATGTGATGG 1033
4573 TAATCGTCAATATAATTTTGTACGTTATGATGATGATGATGATGATGATGATGATG 4632
1034 AGAATTCAGAGAGATATGATCGTCTTAAGGAGGACTTTCAGAGACAGAAAGTTCA 1093
4633 AGAGAACGATAAAGTAGCAATGATTTAGCAAGAAATTAAGAAACAGAAATGTTG 4692
1094 ACATGCTCTCAAAGGACAGGAGTGCACATTTGCTTCTTCCCTTGAAGA -----TA 1147
4693 AGAATTAAGACGATGGGACAAATTAACAAATTTGCTGCTACAAATTAAGAAATTCAA 4752
1148 GCAGCTTCACACTGATGTCGAAATTCGCACTGCTTCCGAGGATGATGATGATGATG 1207
4753 ACCGTGTTGGAACTTTATGATTTGGGAGATGCTTATTAATGAGGAGATGGCAAGTGC 4812
1208 CGGCTTACAAATGCTCCAAAATGCAACACATCACTGTTCTTCTGTTGTTATCAGAG 1267
4813 CTGCTTATCCACTTCTAAAAATTTGAAATTAATCAATCAATCAATCAATCAATCAAT 4872
1268 AGAATTTCTGAGACACTGCTGAGACACTGATGATGATGATGATGATGATGATGATG 1327
4873 CAGATTTGGGATGAAATGAGCAATTTACTATGTTCAAGATGATGATGATGATGATG 4932
1328 AGCTGATGAGTCTTCTGAGAGATTTCAAAA 1362
4933 CACTAATAAGGCTCATATTTCTATTTATCAGGAA 4967

```

```

RESULT 5
US-08-311-731A-130/c
; Sequence 130, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P. C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36941 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
; US-08-311-731A-130

```

```

Query Match 8.8%; Score 133; DB 4; Length 36941;
Best Local Similarity 53.3%; Pred No 2,3e-32;
Matches 511; Conservative 0; Mismatches 425; Indels 23; Gaps 10;
QY 93 GATCCGGAAAACTCGATTCCTAAGGAAGCGGATACAGATCAATCAACGACGAGCTGAT 152
Db 10476 GCTGCCGAGCATCGATGGATCCTGAGGCGGCTACCGCTTATTCCGACTAGCTAAT 10417
QY 153 GCTTGAGGGAATCCAGGTTGAATTTAGCCT--CCTTTGTGACGACATGATGGAGCCT 210
Db 10416 GGTAGACGCGATTTCTCGCTCAACCTAGGCTACTTTTGTCCACCCTGGTGGATCC 10357
QY 211 GAGTGTGATAAATCAATCATGCTCTCCATCAACAAGAACTATGTTGACATGACGAGTAC 270
Db 10356 GAGTTAAGAAATTTGATGGCGGAGACGTTCCGACAAAACATGTTTAGACAAGGACGAT 10287
QY 271 CCCGTCACCCGAACTTCAGAACCGGATGTTGAAACATGATTCACATCTATTCAATGCA 330
Db 10296 CCGG---CGGCCGCGCATCGAACGCGCTGCGTAGCATGCGACCGCTCTCTATTCA 10240
QY 331 CCGTTAGAAAGGCGGAGACCGCTCGGAGTAGAAACCGTTGGATCATCGGAGGCCAPA 390
Db 10239 GCGCTAG--GGACTCAGTACGACAAACCCCTCAATACCTGTGGGTGTTCCATTATCGC 10181
QY 391 ATGTTGCCCGTTGGCTTCAAGCGTAAATGCGACAAAGCGCAAGCTGAAGGCAAA 450
Db 10180 TCCAGGGATGGTGGTGTGATGCTGGTTGCTGCCCAATAAAATTCGCTGGTGGCGTAG 10121
QY 451 CCCCT-----CGATAAACCCCAATTTCCCGGAGCAATGTTCAAGTGTGTGG 502
Db 10120 CTAGTTAGCAGCGTAGCACCCCAATCTGGGATGGCGAGGGTCCAGAGTGTGGTGG 10061
QY 503 AGAAATTCGCTAGTACTTTGAGGTTGAACTTAAAGGAAGTAAATTCAGTGAAGTACT 562
Db 10060 AGAAGTCTCTGCTGTTACTTCCGATGTCGAGC--CGCTATCTGTGATGGAGGAGGATCT 10003
QY 563 ATGTGATGGACCCCTCAACAAGCTGTTGATATG--GTTGATGAGAACACCATTTGTTGTC 620

```

10002 ACCTCATACCTCCGAGCAGGTAATCGAGCCCGTCGACGAGCACTACTTGGTGTGG 9943
 621 GGACATTTCTGT--TCCACTTTAATGGAGAATTCGAAGATGTTAAACTCTTTGAACGAT 678
 9942 TGGCCATCTTTCTGCGCACTTTACCGCGGAACTCGAACCCGTCGCGGAGCTCTGGTC 9883
 679 CWTCTGTTGAAAGAAACAAGAACCGGATGGATACCAATCCACFTGGATGGGCA 738
 9882 GCGCTGGACAAGCTGGACCCGACAAATGGCTTGGACATTTGCAGTGTCCGTCGGTGGCC 9823
 739 AFTGGAGATTCATGTCACCCGTTTTGATCCGGAATGGAAATGGAGACTTTAGACTTCCC 798
 9822 AGCGG-GGTTCCGGATGCCITTCACCGCGAGTTGAGTGGATTCCTGGCTGCC 9764
 799 TFGTGAAGAGATCAATGTF-GAGTGGTCAAGTATGGACT-TGTGTACGCGAGGATG 856
 9763 CGAGTGATGCGATCAATGTAAGTGGCTATAAATATGGGCTCACCCCTACCCCGGTGG 9704
 857 GTTGGTGTATCTGGAGAAACAAGAGGATTTGCCCTGAGAACTCATCTTCATATCAAT 916
 9703 GATTTGGCGTATGGCGGCAACGAATTTGCCGAGGACGTAGTCTTTGGTCACT 9644
 917 ATCTTGGTCTGACCAACCACCTTTACTCTCAATTTCTCCAAAGTTCAAGTCAAGTCA 976
 9643 ACCTGGCGGCGATATGTAGACTTTTCGTAATGAACTTTTCCCGCTGGACAACCAAGTG 9584
 977 TTGCTCAACTACCAACTTATCCGATTTGGCCACGAGGTTACAGAAATGTGATGGAG 1035
 9583 TTAGTCAGTACTCAACTTCTGTGGTGGACGTACGGTACTTGGAAAGTCAATCGAG 9525

RESULT 6
 S-09-056-556-192/c
 Sequence 192, Application US/09056556
 Patent No. 6350456
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Dillon, Davin C.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
 NUMBER OF SEQUENCES: 241
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 FILING DATE: 07-APR-1998
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Maki, David J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 210121.457
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 192:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 359 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-056-556-192

Query Match 4.2%; Score 64; DB 4; Length 359;
 Best Local Similarity 58.3%; Pred. No. 5.9e-11;
 Matches 112; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 QY 844 TACCAGGGATTTGGTGGTGTGATCTGGAGAAACAAGAGGATTTGGCTGAGGAACCTCATC 903
 Db 353 TATCCCGCGTCCGGTTTGTCTGTGGCGGGCCGAGCACCTCCCGGAGGATCTGGTT 294
 QY 904 TTCATATCAATTTCTGTGTGCTGACCAACCCACTTTACTCTCAATTTCTCCAAAGT 963
 Db 293 TTCGGGTCAACTACCTCCGGCGGACATGCCGACCTTCCACCTTGAACCTTCCCGTCCC 234
 QY 964 TCAAGTCAAGTCAATTTGCTCAACTACTCAACTTCTCGGCTGGGGGCGCACCGCTATACC 1023
 Db 233 GGTAAACAGGTTGGGCGCACTACTCAACTTCTCGGCTGGGGGCGCACCGCTATACC 174
 QY 1024 AATGTGATGGAG 1035
 Db 173 AAGGTGATGCAG 162

RESULT 7
 US-09-072-596-187/c
 ; Sequence 187, Application US/09072596
 ; Patent No. 6458366
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Campos-Neto, Antonia
 ; APPLICANT: Houghton, Raymond
 ; APPLICANT: Twardzik, Thomas S.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Hendrickson, Ronald C.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
 ; NUMBER OF SEQUENCES: 350
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/072,596
 ; FILING DATE: 05-MAY-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.417C9
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 187:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 359 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-072-596-187

Query Match 4.2%; Score 64; DB 4; Length 359;
 Best Local Similarity 58.3%; Pred. No. 5.9e-11;
 Matches 112; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 QY 844 TACCAGGGATTTGGTGGTGTGATCTGGAGAAACAAGAGGATTTGGCTGAGGAACCTCATC 903

353 TATCCCGCGTGGGTTGTTGTTGGCGGGCCCGGAGCACTGCCGGAGATCGGTT 294
 904 TTCCATATCAATTAATTTGGTGTGACCAACCCACCTTTACTCTCAATTTCCAAAGGT 963
 293 TTCCGGGTCAACTACCTCGGGGGGACATCCGACCTTACCCTGAATCTTCCCGTCCC 234
 964 TCAAGTCAAGTCATGCTCAATACATACCACTTATCGATTGGCCACGAGGTTACAGA 1023
 233 GGTAACACAGGTGGTGGCCAGTACTACAATCTCTCGGCTGGGGCGGACCGCTATACC 174
 1024 AATGTGATGGAG 1035
 173 AAGTGTGTCAG 162

SULT 8
 -08-232-463-14/c
 Sequence 14, Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHEIFLINGER, F.
 APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOMLOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Foley & Hardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232.463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935.313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 CLONE: IMMEDIATE SOURCE:
 CLONE: PTZPPT-F1s
 US-08-232-463-14

Query Match 2.8%; Score 41.6; DB 1; Length 7218;
 Best Local Similarity 5.0%; Pred.No. 0.01; Mismatches 0; Gaps 0;
 Matches 20; Conservative 208; Indels 172; Indels 0; Gaps 0;
 187 TTGTGACGACATGGATGGGCTGTGATGATAAATCAATCATGTCCTCATCAACAAG 246
 1442 TTGTGACGACATGGATGGGCTGTGATGATAAATCAATCATGTCCTCATCAACAAG 246

Qy 247 AACTATGTTGACATGACGAGTACCCCGTACCACCGAATTCAGAACCGATGTTGAAAC 306
 Db 1382 RRR 1323
 Qy 307 ATGATTCACATCTATTCAATGACCGTTAGAGAGGGGAGACCGCGTCGGAGTAGGA 366
 Db 1322 RRR 1263
 Qy 367 ACCGTTGGATCATCGGAGCCATAATGTTGGCCGTTTGGCCCTTCAAGCGTAATGGCAG 426
 Db 1262 RRR 1203
 Qy 427 AACAAAGCGCAAGCTGAAGCAACCCGTCGATAAACCAACCAATTCACCGGAGCCAA 486
 Db 1202 RRR 1143
 Qy 487 GTTCAAGTGTGTTGGAGAAATTCGCTAGTACTTGGAGTTGAGTTGAACCTAAGGAGTGA 546
 Db 1142 RRR 1083
 Qy 547 TTGAGTGAAGTACTATGTGATGAGCCCTCAACAGCTG 586
 Db 1082 RRR 1043

RESULT 9
 US-09-198-452A-1/c
 ; Sequence 1, Application US/09198452A
 ; Patent No. 659294
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffais, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 1
 ; LENGTH: 1230025
 ; TYPE: DNA
 ; ORGANISM: Chlamydia pneumoniae
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(15000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (15001)..(30000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (30001)..(45000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (45001)..(60000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (60001)..(75000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (75001)..(90000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (90001)..(105000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (105001)..(120000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (120001)..(135000)
 ; OTHER INFORMATION: n=a or c or g or t
 ; NAME/KEY: misc feature
 ; LOCATION: (135001)..(150000)
 ; OTHER INFORMATION: n=a or c or g or t

APPLICANT: Cassell, Gail H.
 APPLICANT: Chen, Ellison Y.
 APPLICANT: Glass, Jennifer S.
 APPLICANT: Glass, John I.
 APPLICANT: Heiner, Cheryl R.
 APPLICANT: Lefkowitz, Elliot
 TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREA PLASMA
 FILE REFERENCE: URB-13452/22
 CURRENT APPLICATION NUMBER: US/09/601,198
 PRIOR FILING DATE: 2000-12-08
 PRIOR APPLICATION NUMBER: 60/073,189
 PRIOR FILING DATE: 1998-01-30
 NUMBER OF SEQ ID NOS: 181
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 60
 LENGTH: 15016
 TYPE: DNA
 ORGANISM: Urea-plasma urealyticum
 3-09-601-198-60

Query Match 2.4%; Score 35.6; DB 4; Length 15016;
 Best Local Similarity 47.7%; Pred. No. 1.6;
 Matches 104; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Y 461 AACCAACATGTCACCGGACCCAAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGACT 520
 b 9593 AACTAAAATTTGCTCAATGATTTTGATGATGTTTTAAAGAAAATCAAACCTTTAAGT 9534

Y 521 TTGAGTTGAACTTAAGGAAGTGAANTGAGTGAAGGATGATGATGAGGACCCCTCAAC 580
 b 9533 TTAATATTCACCAAAATGAAATAATGGTCTAATTTAAATGAGATTGCGCAACATGAAG 9474

Y 581 AAGCTGTTGATATGTTGATGAGAACACCATTGTTGTTGGGACACTTCITGGTCCACTC 640
 b 9473 GTAAAGTAGAAATATTGATGAGGCAAAAATTCITGTTGCTTAATCTTAATAACATTAAG 9414

Y 641 TTAATGGAGAAATCGAAGATGTTAAACTCTGACGAT 678
 b 9413 TTAACAAGAATATAAAGTTTATAAATCTACTTTGAT 9376

RESULT 13
 US-09-026-408-11/5
 Sequence 11, Application US/09026408
 Patent No. 6303338
 GENERAL INFORMATION:
 APPLICANT: Ni et al.
 TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
 TITLE OF INVENTION: INHIBITOR
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVENUE, SUITE 600
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/026,408
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/934,011
 FILING DATE: 15-AUG-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/024,056
 FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.0300002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 515 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 US-09-026-408-11

Query Match 2.2%; Score 33.6; DB 4; Length 515;
 Best Local Similarity 50.0%; Pred. No. 0.78;
 Matches 81; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 525 GGTGAACTTAAAGGAAGTGAATTTGAGTGAAGGATACTATGTGATGACCCCAACAAGC 584
 Db 506 GGTATGTTCAAGNATCCAAACGTCCTTTGGAGTCTACTTTTGGTCTACTTTAATCTGGT 447
 QY 585 TGTGATATGTTGATGAGAACACCATTTGTTGGGACATTTGTTCCGACATTTGGTCCACTTTAA 644
 Db 446 ATTAAAAAAGAAAAGAAAAGTCTTTGAAACACTGACATTTTGTACCAAAATTTT 387
 QY 645 TGGAGAAATTCGAAGATGTTAAACTCTTTGAAACGATCTCTTGGT 686
 Db 386 AATGATGACTTAGAGGATAATATTTGTAANAATCTTTTGT 345

RESULT 14
 US-08-471-033-18
 Sequence 18, Application US/08471033
 Patent No. 5770696
 GENERAL INFORMATION:
 APPLICANT: Warren, Gregory W
 APPLICANT: Kozziel, Michael G
 APPLICANT: Mullins, Martha A
 APPLICANT: Nye, Gordon J
 APPLICANT: Carr, Brian
 APPLICANT: Bessaj, Nalini M
 APPLICANT: Kostichka, N. Kristy
 APPLICANT: Duck, Nicholas B
 APPLICANT: Estruch, Juan J
 TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESS: CIBA-GEIGY Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/471,033
 FILING DATE:
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/314,594
 FILING DATE: 09-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/218,018
 FILING DATE: 23-MAR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/037,057

```

FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: P-40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2004 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2004
OTHER INFORMATION: /note= "Maize optimized DNA
OTHER INFORMATION: sequence for VIPIA(a) 80 kd protein from AB79"
3--08-471-033-18
Query Match 2.2%; Score 33.4; DB 1; Length 2004;
Best Local Similarity 55.7%; Pred. No. 2.2;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
y 1375 CAAGAGAGAGTGAATCTAACACAGCGATACTTGTGTCACGGTGAAGAGCGGATATC 1434
b 1357 CAGGAGAGCTGACACAGACCAGCGACTACTACATCAGCTGTACATGAGAGCGGAGAG 1416
y 1435 GACAAGCAGAGAGATATCATCTACTGGTGGAGAAAGTTTFCGCGCAGAGAGA 1489
b 1417 AACACCCAGTGGAGATCACCATCCGCGGAGATATACCCCATCCACCAAGA 1471
RESULT 15
S-08-471-044-18
Sequence 18, Application US/08471044
Patent No. 5840868
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,044
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIORITY INFORMATION:
APPLICATION NUMBER: US 08/463,483
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLV3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2004 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2004
OTHER INFORMATION: /note= "Maize optimized DNA
OTHER INFORMATION: sequence for VIPIA(a) 80 kd protein from AB78"
US-08-471-044-18
Query Match 2.2%; Score 33.4; DB 2; Length 2004;
Best Local Similarity 55.7%; Pred. No. 2.2;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
Qy 1375 CAAGAGAGAGTGAATCTAACACAGCGATACTTGTGTCACGGTGAAGAGCGGATATC 1434
Db 1357 CAGGAGAGCTGACACAGACCAGCGACTACTACATCAGCTGTACATGAGAGCGGAGAG 1416
Qy 1435 GACAAGCAGAGAGATATCATCTACTGGTGGAGAAAGTTTFCGCGCAGAGAGA 1489
Db 1417 AACACCCAGTGGAGATCACCATCCGCGGAGATATACCCCATCCACCAAGA 1471
Search completed: October 22, 2003, 13:36:58
Job time : 133 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: October 22, 2003, 07:24:21 ; Search time 426 Seconds
(without alignments)
9562.094 Million cell updates/sec

US-10-006-852-1
fect score: 1509
quence: 1 atggtctctcccacgcgft.....agacgagtggtatctgcttaa 1509

ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

arched: 2552756 seqs, 1349719017 residues

tal number of hits satisfying chosen parameters: 5105512

nimum DB seq length: 0
ximum DB seq length: 2000000000
st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- itabase : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with 5 columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 25 rows of sequence alignment results.

Table with 5 columns: ID, AAL43410 standard; cDNA; 1509 BP., 48.8, 1783, 15. Contains sequence alignment data for various species like Arabidopsis thaliana and Listeria monocytogenes.

ALIGNMENTS

Table with 5 columns: RESULT 1, AAL43410, ID, AAL43410 standard; cDNA; 1509 BP., 48.8, 1783, 15. Contains alignment details and organism names.

WPI: 2002-490073/52.
P-PSDB; AK015132.

Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome

Claim 18; Page 53; 63pp; English.

The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the A. thaliana GAD1 coding sequence.

Sequence 1509 BP; 433 A; 321 C; 385 G; 370 T; 0 other;

Query Match 100.0%; Score 1509; DB 24; Length 1509;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Table with 4 columns: ID, Query, Subject, Score. Contains sequence alignment data for various Arabidopsis thaliana sequences (e.g., 1, 61, 121, 181, 241, 301, 361, 421, 481, 541, 601, 661).

Table with 4 columns: Db, QY, ID, Sequence. Contains sequence alignment data for various Arabidopsis thaliana sequences (e.g., 661, 721, 781, 841, 901, 961, 1021, 1081, 1141, 1201, 1261, 1321, 1381, 1441, 1501, 1509 BP).

RESULT 2
ABZ14005
ID ABZ14005 standard; DNA; 1509 BP.
XX
AC ABZ14005;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1810.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX

1501 ATCTGCTAA 1509
 RESULT 3
 AL43415
 AAL43415 standard; cDNA; 1705 BP.
 AAL43415;
 25-SEP-2002 (first entry)
 Tobacco GAD1 coding sequence.
 GAD; plant GABA production regulation; glutamic acid decarboxylase;
 plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
 Nicotiana tabacum.
 Key Location/Qualifiers
 CDS 71..1561
 /tag= a
 /product= "GAD1"
 W0200238736-A2.
 16-MAY-2002.
 07-NOV-2001; 2001WO-US47447.
 07-NOV-2000; 2000US-246367P.
 (EMER-) EMERALD BIOAGRICULTURE CORP.
 Kinnersley AM, Turano FU;
 WPI; 2002-490073/52.
 P-PSDB; RAO15137.
 Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 56-57; 63pp; English.
 The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the tobacco GAD1 coding sequence.
 Sequence 1705 BP; 523 A; 301 C; 385 G; 496 T; 0 other;
 Query Match 58.5%; Score 883.4; DB 24; Length 1705;
 Best Local Similarity 75.2%; Pred. No. 1.5e-260;
 Matches 1125; Conservative 0; Mismatches 351; Indels 21; Gaps 1;
 1 ATGGTGTCTCCCGCGGTATCGGATCGGAGCTCTCCGTCCTCCACTCCCATTCGGCATCA 60
 71 ATGGTGTCTCCAGACAGCGTCGGAAAGTGCCTCCACTCCACTTCGCTCC 130
 61 CGTTACGTCGCTACTCTCCAGTTCCTCCAGTTCGATGCGGAAACTCGATCTCAAGGAA 120
 131 CGATATGTCGCTACTCTCTCCGAGGTTTAAAGATGCCAGAAATTCGATACCAAGGAA 190
 121 GCGGCGTATCAGATCATCAAGCAGCGTGTCTTGCAGGGAATCCACGCTTCACTTA 180
 191 GCAGCATATCAAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAAGACTAAATTTA 250

181 GCCTCCTTTGTGACGACATGGATGGAGCCCTGAGTGTGATAAACTCATATGCTCCTCCATC 240
 251 GCATCTTTTGTGACCAACATGGATGGAACACAGAGTGTAAACAACATGATGATGGATTCGATT 310
 241 AACAGAACTATGTTGACATGGAGCGAGTACCCTGTCACCCAGCACTTCCAGAACCTTCAAGACCGATGT 300
 311 AACAGAAATACGTTGACATGGATGAAATCCCTGTAAACCACTGAACTTCAAGATTCGATGT 370
 301 GTGAACATGATGGACATCTATTCAATGCACCGTTAGAAGAGCGGAGACCGCGCTGGGA 360
 371 GTAAACATGATGACTATTGTTAAACGACCCACTTGGAGATGGAGAGACTGCAGTTGGA 430
 361 GTAGAACCTGTTGATCATCGGAGCCATAATGTTGGCCGGTGTGGCCCTTCAAGCGTAA 420
 431 GTTGAACCTGTTGATCCTCTGAGCTATTATGCTGCTGATGCTTGTGCTTCAAGAGAAA 490
 421 TGGCAGAACACGCGCAAGCTGAAAGCCAAACCCGTCGATAAACCCCAACATTTGTCACCGGA 480
 491 TGGCAAAATAAATGAAAGCCCAAGGCCCTGTGCACAAAGCCCAATAATGTCACCTGGT 550
 481 GCCAATGTTCAAAGTGTGTTGGGAGAAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
 551 GCCAATGTCAGGTGTGTTGGGAGAAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610
 541 GTGAAATGAGTGAAGGATCTATGATGAGACCCCTCAACAGCTGTGATATGTTGAT 600
 611 GTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
 601 GAGAACACCAATTTGTTGGGACATTTCTTGGTTCCTACTCTTAAATGGAGAAATTCGAAAGAT 660
 671 GAGAACACCAATTTGTTGGGAGATCTCTTGGTTCCTACTCTTAAATGGAGAAATTTGAAAGAT 730
 661 GTTAAACTTTGAAACGATCTTCTTGGTTCGAAAAGAAACAAAGAACCCGATGGGATACCA 720
 731 GTTAAAGCTTTGAAATGACCTCTTGTGATGAGAAAGAAACAAAGAACCCGATGGGACTCCA 790
 721 ATCCACGTTGAGTGGGCAAGTGGAGGATTCATTCACCGTTTTGTTGATCCGGAAATCGAA 780
 791 ATTCAATGTTGAGTGGGCAAGTGGAGGATTCATTCACCGTTTTGTTGATCCGGAAATCGAA 850
 781 TGGGACTTTAGACTTCCCTTGGTGAAGATATCAATGTGAGTGGTCAAGATGAGACTT 840
 851 TGGGACTTTAGACTTCCCTTGGTGAAGATATCAATGTGAGTGGTCAAGATGAGACTT 910
 841 GTGTACGAGGATGTTGGTGTATCTGGAGAAACAAAGAGATTTGCCCTGAGAACTC 900
 911 GTTATGCTGTTGTTGGGCAATTTGGAGGAAATAGGAAAGACTTACCTGAGAACTT 970
 901 ATCTTCCATATCAATATCTTGGTGTGACCAACCCACTTTACTCTCAATTTCTCCAAA 960
 971 ATCTTCCATATTAATCTTGGTGTGATCAACTACTTTCACTCTCACTTCTCTAAA 1030
 961 GGTTCAGTCAAGTCAATGCTCAATFACCAACTTATCCGATTTGGCCACGAGGGTTAC 1020
 1031 GGTTCAGTCAAGTCAATGCTCAATFACCAACTTATCCGATTTGGCCACGAGGGTTAC 1090
 1021 AGAATGTGATGGAGAAATTCAGAGAGATATGATCGTCTTAAGGAAAGACTTTGAGAAG 1080
 1091 AAGAATGTGATGGAGAAATTCAGAGAAATTCAGAGGATTCAGAGAAAGACTTTGAAA 1150
 1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTCCACTTGTCCCTTTCTCCTTG 1140
 1151 AGTGAAGATTCACATAATAATCCAAAGAAATGGAGTTCCTATTAGTAGCTTTCTCTCT 1210
 1141 AAAGATGACAGTGTACACTGAGTTCGAAATCTCCGACATGCTTCCAGGTATGGATGG 1200
 1211 AAAGAACAGTGTACACAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1270
 1201 ATAGTCCGCGCTACCAATGCTCCAAATGCTCCAAATGCTCCAAATGCTCCAAATGCTCC 1260
 1271 ATTATCTTGCATATCTATGCTCCAAATGCTCCAAATGCTCCAAATGCTCCAAATGCTCC 1330
 1261 ATCAGAGAAATTTCTCGAGAAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGT 1320

1331 ATTAGAGAGATTTCCTCCGTACACTCCGCGAGCGACTGGTAATAGACATGGAABAAGTC 1390
 1321 ATGCGTGAAGCTCGATGAGCTTCCCTTCGAGAGTGAATTCACAAAATATACACTTGGACAAAGAG 1380
 1391 CTCACGAGCTAGACACACTTCCGCGAGGGTCAACGCTAAGCTAGCTAGCCGTCGCGAGGCG 1450
 1381 AAGAGTGAATCTAACAGGATPAACTTGTGATGATCAGGTGAAGAGAGCGATATCGACAAAG 1440
 1451 AATGG-----CAGCGCGTGCATPAAGAAAACAGATAGAGAGTG 1489
 1441 CAGAGAGATATCACTGCGTGAAGAAGTGTGCGCCGACAGGAGAGAGCGAGT 1497
 1490 CAGCTTGAGATTACTACTGATGGAAGAAATTTGTTGCTGATGAAGAAGAAGACT 1546

RESULT 4
 SL45417
 D AAL43417 standard; DNA; 1784 BP.
 X AAL43417;
 X T 25-SEP-2002 (first entry)
 X E Petunia GAD coding sequence.
 X X GAD; plant GABA production regulation; glutamic acid decarboxylase;
 W plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
 S Petunia sp.
 X X WO200238736-A2.
 X N 16-MAY-2002.
 X D 07-NOV-2001; 2001WO-US47447.
 X F 07-NOV-2000; 2000US-246367P.
 X R (EMER-) EMERALD BIOAGRICULTURE CORP.
 X A Kinnersley AM, Turano FJ;
 X J WPI: 2002-490073/52.
 X R P-PSDB; AAO15139.

Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 58; 63pp; English.
 The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the petunia GAD coding sequence.
 Sequence 1784 BP; 550 A; 298 C; 400 G; 536 T; 0 other;
 Query Match 57.7%; Score 871.4; DB 24; Length 1784;
 Best Local Similarity 74.9%; Pred. No. 7.4e-257;
 Matches 1121; Conservative 0; Mismatches 366; Indels 10; Gaps 2;

1 ATGGTGTCTCCACGCGGTATFCGGAGTGGAGCGCTCCGTCACCTCCACATCGCATCA 60
 72 ATGGTGTCTCAAGACAGTGTCCGAGAGCGATGTGTCATTCCTCCAGCTTGTCTTCT 131

QY 61 CGTTACCTCGTACTTCACTTCTAGGTTCAAGATCCCGAAAACTCGATTCTTAAGGAA 120
 Db 132 CGATATGTTGCACTTCTTCCAGGTTAAAATCCAGATAAATTCGATACAAAAGAA 191
 QY 121 GCGCGTATCAGATCAACAGACGAGCTGAGCTTGGACGGGAATCCACGGTTGAACCTTA 180
 Db 192 GCACATATCAGATCAAAAATGATGAATGATGATGATGATGATGATGATGATGATGATG 251
 QY 181 GCCTCCCTTTGTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 Db 252 GCTTCTTTTGTTCACACUUGATGGAACCCAGAGTGTGATGATGATGATGATGATGATG 311
 QY 241 AACAAAGACTATGTTGACATGAGTACCCCGTACCACCGAACTTCAGAACCGAGT 300
 Db 312 AACAAAGACTATGTTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 371
 QY 301 GTGAACATGATGACATCTTCAATGACACCGTGTAGAGAGCGGAGACCCGCTCGGA 360
 Db 372 GTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 431
 QY 361 GTAGAAACCGTTGGATCATCGGAGCCATAAATGTTGCCGGTTTGGCCCTTCAAGCGTAAA 420
 Db 432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTCTGATTTAGCTTTCAGAGAAA 491
 QY 421 TSSCAGAACAGCCCAAGCTGAAGCGAAACCCGTGATAAACCCAACTTGTCAACCGA 480
 Db 492 TGGCAGAACAAAATGAAAAGCCCAAGGCAACCCCTGTGACAAGCCCAACATGTTACTGT 551
 QY 481 GCCAATGTTCAAAGTGTGTTGGAGAAATTCGTAGTACTTTGAGGTTGAACTTAAGGAA 540
 Db 552 GCAATGTTCAAGTGTGCTGGGAAATTTGCAAGGTATTTTGAAGTGGAGCTTAAAGGAA 611
 QY 541 GTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 Db 612 GTAAAGCTTGTGAAGGATATGATGATGATGATGATGATGATGATGATGATGATGATG 671
 QY 601 GAGAACCACTTGTGTTGGGCACTTCTTGGTTCCACTCTTAATGGAGAAATTCGAAAGAT 660
 Db 672 GAACACCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 731
 QY 661 GTTAAACTTTGAAACGATCTTTGGTCCGAAAAAGAAACAAAGAAACCCGATGGGATACCCA 720
 Db 732 GTTAAAGCCTTGAATGATCTTTGGTCCGAAAGAAACAAAGAAACCCGATGGGATACCCA 791
 QY 721 ATCCACGTGATGCGGCAAGTGGAGGATTCATGTCACCGTTTTTGTATCGGAAATGGA 780
 Db 792 ATTATGTTGATGAGCAAGTGGTGGATTTATGACCCGTTCAATACCAGAGCTTGGAG 851
 QY 781 TGGCACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGTCAAGATGAGACTT 840
 Db 852 TGGCACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGTCAAGATGAGACTT 911
 QY 841 GTGTACCAGGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
 Db 912 GTCTATGCTGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 970
 QY 901 ATCTTCCATATCAATATCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
 Db 971 ATCTTCCATATCAATATCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1030
 QY 961 GGTTCAAGTCAAGTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAAT 1020
 Db 1031 GGTTCTAGCCAAAGTAAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCA 1090
 QY 1021 AGAATGATGAGGAAATTCAGAGAGAAATGATGATGATGATGATGATGATGATGATGATG 1080
 Db 1091 AAGAATGATGAGGAAATTCAGAGAAATGATGATGATGATGATGATGATGATGATGATG 1150
 QY 1081 ACAGAAAGGTTCAACATGCTTCAAGGACGAGGAGTGCACCTTGTGCTTTCCTTTG 1140
 Db 1151 ACAGAAAGTCAACATAAATCTCCAAAGAAATGAGTACCTTTAGTAGCATTTCTCTCT 1210
 QY 1141 AAGATAGCAGCTGCACATGAGTTGAAATCTCCGACATGCTCCGAGGATGGATGG 1200

1211 AAAGCAACAGGCAACACACAGGTTCCGAGATTTCTGAAATTTAAGGAGATTTGGTTGG 1270
 1201 ATAGTGGCGGCTACACATGCTCCAAATGCACAAACACATCACTGTTCTTCGTGGTT 1260
 1271 ATTGTTCTGATATACATGCTCCCAACCAACGCAACACATTAACAGTTCCAGAGTTGG 1330
 1261 ATCAGAGAGATTTCTCGAAGACACTCCCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 1331 ATCAGAGAGATTTCTCCGTAAGCTTGCAGAGACTGGTAGAGACATCGAAGAGTIC 1390
 1321 ATGCGTAGCTCGTAGAGCTTCCCTGAGAGTGAATTCACAAAATATCACTTTGACACAGAG 1380
 1391 CTTATGAACTTGACACACTCCCTGCACTGCTCAATGCTAAGCTGCTGCTGCTGCGGAGAG 1450
 1381 AAGAGTGAATCTAACAGGATTAACCTTGTATGCTCAAGCTGAGAGAGCGGATATCGACAAG 1440
 1451 CAGCGGGTGGGATGGC-----AGCGAGTGCATTAAGAAAACAGATAGCGGAGTGS 1501
 1441 CAGAGAGATATCATCTGCTGGAAGAGTTTGTCCCGCACAGGAGAGACCGAGT 1497
 1502 CAGTTGGAGATGATAACTGCATGGAGAGATTTCTTGAAGAAAAGAGAAAGAACT 1558

RESULT 5
 AL43411
 D AAL43411 standard; cDNA; 1665 BP.
 X
 C AAL43411;
 X
 T 25-SEP-2002 (first entry)
 X
 E A thaliana GAD2 coding sequence.
 W
 X GAD; plant GABA production regulation; glutamic acid decarboxylase;
 W plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
 W Arabidopsis thaliana.
 S

X Key Location/Qualifiers
 H 17..1501
 T CDS /*tag= a
 T /*product= "GAD2"
 X
 X WO200238736-A2.
 X
 X 16-MAY-2002.
 X
 X 07-NOV-2001; 2001WO-US47447.
 X
 X 07-NOV-2000; 2000US-246367P.
 X
 X (EMER-) EMERALD BIOAGRICULTURE CORP.
 X
 X Kinnersley AM, Turano FJ;
 X
 X WPI; 2002-490073/52.
 X
 X P-PSDB; AAO15133.

Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 53-54; 63pp; English.
 The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced

CC ability to tolerate environmental or other stresses. The present sequence
 CC is the A. thaliana GAD2 coding sequence.
 XX

SQ Sequence 1665 BP; 490 A; 311 C; 423 G; 441 T; 0 other;
 Query Match 57.78; Score 870.2; DB 24; Length 1665;
 Best Local Similarity 75.18; Pred. No. 1.7e-256;
 Matches 1110; Conservative 0; Mismatches 348; Indels 21; Gaps 1;

QY 31 GACGTCCTCCCTCCACTCCACATTCGCATCCGATCAGCTTACGTCGGTACTTCACTTCCCTAGTTC 90
 Db 44 GATGAATCTGTCGCACCATGTCGGATCTCGCTATGTTCCGACTACACTTCCCAAGTAT 103
 QY 91 AAGATGCCGAAAACCTCGATTCCTAAGCAAGCGCGCTATCAGATCATCAACGACGAGCTG 150
 Db 104 GAGATTTGTGAGAAATTCGATACCGAAGACGCTGCATATCAGATCATAAAGATGAGCTG 163
 QY 151 ATGCTTGACGGGAATTCACCGTTCGAACTTAGCCCTCCTTTGTGACGACATGGATGGAGCC 210
 Db 164 ATGCTTGATGGTAAACCGGAGCTTAACTTAGCTTGGTTTGTGACTACATGGATGGAACCA 223
 QY 211 GAGTGTGATAAACTCATCATGCTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
 Db 224 GAGTGTGACAAACTCATCATGCTCCTCATCAAGAACTACCTGTTGATATGATGAGTAC 283
 QY 271 CCCGTCACCCGAACTTCAGAACCCGATGTTGAAACATGATCCACATCATTTCAATGCA 330
 Db 284 CCTGTCAAACTGAGCTCCAGAACCCGATGTTGAAACATTTAGCTCGACTGTTCAATGG 343
 QY 331 CCGTTAGAAGAGGGGAGACCGCGTCCGAGTAGGAAACCGTTCGATCATCGGAGGCCATA 390
 Db 344 CCACTCGAGGAATCTGAGACCGCGTGGAGTAGGACAGTTGGTCTTCAGAAGCCATC 403
 QY 391 ATGTTGGCCGTTTGGCTTCAAGCGTAAATGTCAGAAACGCGCAAAAGCTGAAGGCAAA 450
 Db 404 ATGTTAGCCGATTTGGCTTCAAAAGAAATGGCAGAAACAACGCAAGGCTGAGSGTAA 463
 QY 451 CCCGTGATTAACCCAACTATGTCACCGGAGCCAAATGTTCAAGTGTGTTGGGAGAAATTC 510
 Db 464 CCCTATGACAAAACCCAACTTGTCACTGGAGCCAAATGTTCAAGTTTCTGGGAGAAATTC 523
 QY 511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAATTTGAGTGAAGGATACTATGTGATG 570
 Db 524 GCTCGTACTTTCGAGTGGAGCTTAAGGAAGTAAACCTAAGTGAAGTACTACTAGTGTATG 583
 QY 571 GACCTTCAAGAGCTGTTGATGTTGATGAGAACACCAATTTGTTGTTGTTGCGGACATTC 630
 Db 584 GATCCAGACAAAGAGCAGGAAATGTTAGACGAGAAACAACAATCTGTGTCCGACCATATTG 643
 QY 631 GGTTCCTCTTAATGGAGAAATTCGAAAGATGTTAAACTTTGAAACGATCTCTTTGGTCGAA 690
 Db 644 GGATCCACTCAACGGTGAATTCGAAAGAGCTGAAAGCTCTCAATGACTTGTAGTCAAG 703
 QY 691 AAGAACAAAGAAACCCGATGGATFACACCAATCCAGTGGATCCGCAAGTGGAGGATTC 750
 Db 704 AAAAAGCAGGAGACTGGTTGGAAACACACACCCAGTCCAGCTGGATGACGCAACTGGAGGTT 763
 QY 751 ATTGCACCGTTTGTATCCGGAAATTCGAAATGGACTTGAAGTTCCTTCCCTGTTGAGAGT 810
 Db 764 ATAGTCCGTTTATCTATPCTGAAATTAAGATGGACTTTAGACTTCTCTTTGGTTAAGAGT 823
 QY 811 ATCAATGTGAGTGTTCACAAGTATGGACTTGTGTACGACGGGATTTGGTGGTGACTCTGG 870
 Db 824 ATCAACGTGAGTGTTCACAAGTATGGACTGTTCTATGCTGTTGTTGTTGGTCTGTTG 883
 QY 871 AGAAACAAGAGGATTTCCCTGAGGAACTCATCTTCCATATCAATTAATCTTGGTCTGAC 930
 Db 884 AGGCAGCAGAGGATTTCCCTGAGGAGCTTATCTTTTCAATTAATTAATCTTGGTCTGAT 943
 QY 931 CAACCCACTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCAATTTGCTCAACTACTAC 990
 Db 944 CAACCCACTTTACTCTCAATTTCTCCAAAGGATTCAGCCAAATTAATTTGCTCAACTACTAC 1003

991 CAACCTATCCGATTTGGCCACGAGGGTTACAGAAATGTGATGGAGAAATTCAGAGAGAAAT 1050
 1004 CAGCTCATCTCTTGGATTGGAGGGTACAAAATGTGATGGAGAAATTCAGAGAGAAAT 1063
 1051 ATGATCGTCTAAGGGAAGACTTGAAGACACAGAAAGTTTCAACATCGTCTCAAAAGGAC 1110
 1064 ATGTTGGTCTCAAGAAGGGATGAGAAAACAGAGCGTTTCAACATAGTCTCAAAAGGAC 1123
 1111 GAGGAGTGGCACTTGTGCTTCTCCCTTGAAGATAGCAGCTGTCACTGATGTTGCGAA 1170
 1124 CAAGAGTGGCAGTCTGATGCTTCTCAAGACCATAGTTCACAACGAGTTCGAG 1183
 1171 ATCTCCGACATGCTTCCAGGATGATGAGATGAGTGGCGGCTTACACAATGCTCCAAAT 1230
 1184 ATCTCTGAGATGCTACGTCGTTTGGCTGGATCGTCCAGCTTACACTATGCTCCCGGAT 1243
 1231 GCACAAACACATCACTGTTCTTCGTTGTTTATCAGAGAAGATTTCTCGAAGACATCGCT 1290
 1244 CCACAGCACATCAGGTTCTCGGTTGTTCTACAGGAAGATTTCTCAGAACACTCGCG 1303
 1291 GAGACATTTGATCGATATAGAAAAGTATGATCGTGGAGCTTCCCTTCGAGA 1350
 1304 GAGACATTTGATGATATTTCAAGTGTCTTCAAGCTAGATACCTTGCCTTCCAAG 1363
 1351 GTGATTCACAAAATACACTTGGCAAGAGAGTGAATCTAACAGCGATAACTTCAATG 1410
 1364 AT-----ATCTAAGAAGATGGAATGAAAGGATTCGCGGAAAAT 1402
 1411 CTCACGTTGAAGAAGAGCGATATCGACAAGCAGAGAGATATCATCTGGCTGGAAGAAG 1470
 1403 GTAAAGGAGAGAAGATGAGAAAGGATTTCTGATGGAAGTATTTGATGAGAGAG 1462
 1471 TTTGTCGCCGACAGGAAGACGAGTGGTATCTGCTAA 1509
 1463 TTTGTGAAGGAGGAAAGAAATGAATGGTGTGTGCTAA 1501

RESULT 6
 AL43416
 D AAL43416 standard; cDNA; 1771 Bp.
 X C AAL43416;
 X C
 X T 25-SEP-2002 (first entry)
 X E Tobacco GAD2 coding sequence.
 X X
 W J GAD; plant GABA production regulation; glutamic acid decarboxylase;
 W W plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
 X X Nicotiana tabacum.
 X S
 X X Key Location/Qualifiers
 X H 67..1157
 X T CDS /*tag= a
 /product= "GAD2"
 W0200238736-A2.
 16-MAY-2002.
 07-NOV-2001; 2001WO-US47447.
 XX
 XX 07-NOV-2000; 2000US-246367P.
 XX
 XX (EMER-) EMERALD BIOAGRICULTURE CORP.
 XX
 XX Kinnersley AM, Turano FJ;
 XX PI
 XX WPI; 2002-490073/52.
 XX DR P-PSDB; AAL015138.
 XX
 XX Making transformed plants that selectively increase gamma-aminobutyric

acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 57; 63pp; English.
 The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the tobacco GAD2 coding sequence.
 Sequence 1771 Bp; 548 A; 295 C; 393 G; 535 T; 0 other;
 Query Match 56.8%; Score 857.8; DB 24; Length 1771;
 Best Local Similarity 74.1%; Pred. No. 1.1e-252;
 Matches 1109; Conservative 0; Mismatches 367; Indels 21; Gaps 1;
 QY 1 ATGTFGCTCTCCACGCGGATCGAGTCCGAGTCTCCGTCACCTCCACATTCGATCA 60
 Db 67 ATGGTTCGTCRAAGACAGCGTCCGAAAGTGAAGTCTCCGTTCCACCTTCGCTCC 126
 QY 61 CGTTACGTCGTTACTTCACTTCCCTAGGTTCAAGATCCGGAAAACCTCGATTCCTAAGGAA 120
 Db 127 CGATATGTTCCAACTTCTTCCAGGTTTAAATGCCAGAGAAATCAATACCAAGGAA 186
 QY 121 GCGGCTATCAGATCATCAAGCGAGCTGATGCTTGCACCGGAAATCCACGGTTGAAC 190
 Db 187 GCAGCATATCAGATTTAAATGATGAGCTTATGTTAGATGGAATCCAGGCTAAATTA 246
 QY 181 GCCTCCTTTGTCAGACATGATGGAGCTGAGTGTGATAACTCATCATGTCTCCATC 240
 Db 247 GCATCTTTCGTTACAACAATGGATGGAGCCAGAAATGTAATACGTTAATGATGAT 306
 QY 241 AACAGAACTATGTTGATGAGAGAGTACCCCGTCAACCCGAACTTCAGAACCGATGT 300
 Db 307 AACAGAACTACGTTGATGATGAAATACCCTGTAAACCACTGAGCTTCAGAAATCGAT 366
 QY 301 GTGAACATGATTCACATCTATCAATGCACCGTTAGAAAGGCGGAGACCGCGTCGGA 360
 Db 367 GTAATATGATGATCTATTTGTTATGACCACTTGSAGATGAGAGACTGCAGTTGGA 426
 QY 361 GTAGAACCGTTGATCATCGGAGCCATAATGTTGGCCGTTTGGCTTCAAGCGTAAA 420
 Db 427 GTTGGAACTGTTGGATCCTCTGAAGCTATTTATGCTTGTGGATTAGCTTTAAGAGAAA 486
 QY 421 TGGCAGAAACGCGAAAGCTGAAAGCAACCCGTCGATAAACCCCAACTTGTCCCGGA 480
 Db 487 TGGCAAAATAAATGAAAGCCCAAGGCAAGCCCTTTGATGAAAGCCCAATATGTCACCGGT 546
 QY 481 GCCAATGTTCAAGTGTGTTGGAGAAATTCGATAGTACTTTTGGAGTGAACCTTAAGGAA 540
 Db 547 GCTAATGTCAGGTTGTTGGGAAATTTGCAAGTATTTTGAAGTGGAGTTGAAAGAA 606
 QY 541 GTGAATGAGTGAAGATATCTGATGGAACCCCAAGCTGTTGATGATGTTGAT 600
 Db 607 GTAAAATGATGATGATATCTATGATGGAACCCCTGAGAAAGCTGTGAAATGTTGGAT 666
 QY 601 GAGAACACCAATTTGTTGGGCAATTTCTTGGTTCCACTCTTAATGGAGAAATTCGAAAT 660
 Db 667 GAGAAATACCAATTTGTTGCTGCTTACTTAGTTCACACTCAATGTTGAATTTGAGAT 726
 QY 661 GTTAAACTTTGAACGATCTCTTGGTCGAAAAAACAAGAAACCCGATGGGATACCA 720
 Db 727 GTTAAAGCGTTTGAATGACCTTTTGAATGAGAGAAACAAGAAACCCGGTGGACACTCCA 786
 QY 721 ATCCACGTTGATGGCAAGTGGAGGATTCATTTGCAAGCTTTTGTATCCGGAATGGAA 780
 Db 787 ATTCAATGTTGATGAGCAAGTGGTGGATTTATTTGCACCAATTCCTTTTACAGAGCTTGAA 846

781 TGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGAGTGGTCCACAAGTATGGACTT 840
 847 TGGGACTTTAGATTGCCATTTGGAGAGATTAATGAGTGGTCCAAAATATGTTCTT 906
 841 GTGTACCCAGGGATGTTGGTGGTGAATCGAGAAAACAAGAGGATTTGCCTAGGAACATC 900
 907 GTCTATGCTGATGTTGGTGGGCAATTTGGAGGAATAAGGAAGACTTGGCTGATGAACAT 966
 901 ATCTTCAATCAATATCTGTTGGTGGTGAACCAACCCACCTTTACTCTCAATTTCCCAA 960
 967 ATTTTCCACATCAATTAACCTTGGTGGTGAATCAACCTTCTCACTTCTCACTTCTCAA 1026
 961 GGTTCAGAGTCAAGTCAATGCTCAATATACCACTTATCCGATTTGGCCACAGGGTTAC 1020
 1027 GGTTCAGCAAGTAAATGCTCAATATACCACTTATTCGCTTGGTGGTGGTGGTGGTGG 1086
 1021 AGAATGTGATGGAGATTCAGAGAGAAATATGATCGTCTTAAGGAAGGACTTGAAG 1080
 1087 AAGAATGTTATGGAGAAATTCAGAAATGCAAGGATTAAGAGAAAGAAATTTGAAAA 1146
 1081 ACAGAAAGGTTCAACATCGTCTCAAGCAGGAGGAGTCCACTTGTGGCTTCTCTCTTG 1140
 1147 AGTGAAGATTCACATAATCTCCAAAGAAATGGAGTTCCTTAGTAGCAATTTCTCTT 1206
 1141 AAGATAGAGCTGTCACTGATGTTCCAAATCTCCGACATGTTCCAGGTATGATGG 1200
 1207 AAGACAAACAGTCAACCAATGATGTTCCAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1266
 1201 ATAGTGGCCCTACCAATGCTCCAAATGCAACACATCACTTCTTCTGTGTGTT 1260
 1267 ATTGTTCTGGCAATATATGCTCCAAATGCTCAACATGTCACAGTTCACAGATTGTC 1326
 1261 ATCAGAAAGATTTCTCGAAGAACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 1327 ATTAGAGAAATTTCTCCGACACTAGCGAGGACTGTTAATAGCAATTTGAAAAGTC 1386
 1321 ATGGGTAGCTGATGATGTTCTCGAGAGTGAATTCACAAAATATCACTTTGACAAAGAG 1380
 1387 TTCCACGGAGTAGACACTTCCGGGAGGCTCAACGCTAAGCTAGCCGTGGCCGAGGGC 1446
 1381 AAGAGTGAATCTACAGGATACTTGTGTTCCAGCTGATGATGATGATGATGATGATGAT 1440
 1447 AATGG-----CAGCGGCTGCATTAAGAAAACAGATAGAGAAGTG 1485
 1441 CAGAGATATCATCACTGGTGGAGAAATTTGTCGCGCACAGGAGAGAGAGAGT 1497
 1486 CAGCTAGAGATTACTACTGCTGATGTTGAAATTTGTTGCTGATAGAGAAAGAGACT 1542

XX PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Harper JF, Kreps J, Wang X, Zhu T;
 PI MPI; 2002-304127/34.
 XX
 DR Identifying a stress condition to which a plant cell has been exposed
 XX and producing plants with increased tolerance to these abiotic stresses
 PT
 PT Claim 144; SEQ ID NO 937; 57pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising:
 CC (a) contacting nucleic acid representative of expressed polynucleotides
 CC in the plant cell with an array of probes representative of the plant
 CC cell genome; and
 CC (b) detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.
 XX SQ Sequence 1482 BP; 402 A; 331 C; 378 G; 371 T; 0 other;

Query Match 56.6%; Score 854.2; DB 24; Length 1482;
 Best Local Similarity 76.0%; Pred. No. 1.3e-251;
 Matches 1054; Conservative 0; Mismatches 333; Indels 0; Gaps 0;
 QY 1 ATGGTGTCTCCCGCGGTATCGAGTGGAGTCCCGTCCACTCCATTCGCACTCA 60
 Db 1 ATGGTTTTCTTAAGACAGTTTCCGAACTGATGATCTCAATCCATTCACATTTGCTTCT 60
 QY 61 CGTTACGTCCTTCACTTCCCTTAGGTTCAAGTCCGGAAAACCTCGATTCCTAAGGAA 120
 Db 61 CGTTACGTCCTTCCAGGATTCGAAATGCTGAGAACTCAATCCCAAGAA 120
 QY 121 GCGGGTTCAGATCAACGACGAGCTGATGCTTACCGGAAATCCACGGTTGAACCTTA 180
 Db 121 GCAGTTTACCAATATCAACGACGAGCTAATGCTGATGTAACCAAGCTGAACCTA 180
 QY 181 GCCTCTTTGTGACGACATGATGGAGCTGAGTGTGATAACTCATCTCCTCCATC 240
 Db 181 GCTTCTTCTGACCAATGATGGCCAGAAATGTGACAACTCATGATGGAGTCCATC 240
 QY 241 AACAAAGACTATGTTGATCATGGACGAGTACCCTGTACCACCGAACTTCAGAACCGATGT 300
 Db 241 AACAAAGACTATGTTGATCATGGACGAGTACCCTGTACCACCGAACTTCAGAACCGATGT 300
 QY 301 GTGAACATGATGACATCTATCAATGACCGCTTAGAAGAGGCGGAGACCGCGTCCGA 360
 Db 301 GTTAACATGATGACGCTCTTCAACGCGCGCTTGTGACGGTGAAGTGCCTGTTGT 360
 QY 361 GTAGGAACCGTTGATCATCGAGCCCAATATGTTGGCCGTTTGGCTTCAAGCGTAAA 420
 Db 361 GTTGGCACCGTGGATCGTGGAGGCGATTAATGTTGGCCGTTTGGCTTTAAGAGCAA 420
 QY 421 TGGCAAAACAGCGCAAGCTGAAAGCAACCGCTGATTAACCAACCAACATGTCACCGGA 480
 Db 421 TGGCAAAATAAGCGTAAAGCCCAAGGGCTTCTTATGATTAAGCCCAATATGCTAACCGGT 480
 QY 481 GCCAATGTTCAAGTGTGGAGAAATTCGCTAGGTACTTTGAGTTGAACTTAAAGGAA 540
 Db 481 GCTAAATGCTCAGGTTTCTGGAGAAATTCGCAAGGATTTCCGAAGTGGAGCTTAAAGAA 540
 QY 541 GTGAAATGAGTGAAGGATATGATGATGGACCTCAACAAAGCTGTTGATGATGTTGAT 600

RESULT 7
 ID ABZ13132
 ID ABZ13132 standard; DNA; 1482 BP.
 AC ABZ13132;
 AC
 DT 21-JAN-2003 (first entry)
 DE
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 937.
 DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 DE Arabidopsis thaliana.
 DS
 DS WO200216655-A2.
 DS 28-FEB-2002.
 DS 24-AUG-2001; 2001WO-US26685.
 DS 24-AUG-2000; 2000US-227866P.
 DS 26-JAN-2001; 2001US-264647P.
 DS 22-JUN-2001; 2001US-300111P.

```

541 GTGAACCTAAGAGAACTATTACGTGATGGACCCTGTAAAGCGGTCGAAATGTTAGAC 600
601 GAGAACACCACTTTGTTGGGACATTTCTTGGTTCCACTCTTAATGAGAAATTCGAAAGAT 660
602 GAAACACAAATTTGTTGGTCCCACTCTCGTTCACGTTAACCGTGAATTCGAAAGAC 660
661 GTTAAACTTTGAAACGATCTCTTGGTCAAAAGAAACAAAGAAACCGGATGGGATACCA 720
662 GTTAAAGTCTCTCAACGACCTCTTGTGAGAAAACAAAGAAACCGGATGGGACGCCA 720
721 ATCCACGCGGATCGGCAAGTGGAGGATTCATCCACCGCTTTTGTATCCGGAAATGGA 780
722 ATACACGTTGACGCGGATCGGCAAGTGGAGGATTCATCCACCGCTTTTGTATCCGG 780
781 TGGGACTTTAGACTTCCCTTGGTGAAGAGATCAATGAGTGGTCAAAAGTGGGATG 840
782 TGGGATTTCCGCTACCGTGGTTAAGAGATTAATGAGTGGTCAAAAGTGGGATG 840
841 GTGTACCGAGGATTTGGTGGTGAATCTGGAGAAAACAAAGAGATTTGCCGTGAGGAA 900
842 GTTAAACCGGATTTGGTGGTGAATCTGGAGAAAACAAAGAGATTTGCCGTGAGGAA 900
901 ATCTTCCATATCAATTAATCTTGGTGGTGAATCTGGAGAAAACAAAGAGATTTGCC 960
902 ATCTTCCATATCAATTAATCTTGGTGGTGAATCTGGAGAAAACAAAGAGATTTGCC 960
961 GGTTCAGTCAAGTCAATTAATCTTGGTGGTGAATCTGGAGAAAACAAAGAGATTTGCC 1020
962 GGTTCAGTCAAGTCAATTAATCTTGGTGGTGAATCTGGAGAAAACAAAGAGATTTGCC 1020
1021 AGAATCTGATGGAGAAATTCAGAGAAATATGATCGTCTTAAAGGAGGACTTTGAGA 1080
1022 CGCAATGTGATGGAGAAATTCAGAGAAATATGATCGTCTTAAAGGAGGACTTTGAGA 1080
1081 ACAGAAAGTTTCAACATCGTCTTAAAGGAGGAGGAGTGCACCTTTTCGCTTTCTCT 1140
1082 ACAGAAAGTTTCAACATCGTCTTAAAGGAGGAGGAGTGCACCTTTTCGCTTTCTCT 1140
1141 AAAGATGACGCTGTACACTGAGTTCGAAATCTCCGAAATTCGCTTTTCGCTTTCTCT 1200
1142 AAAGATGACGCTGTACACTGAGTTCGAAATCTCCGAAATTCGCTTTTCGCTTTCTCT 1200
1201 ATAGTGGCGCTACAGAAATTCGAAATTCGCTTTTCGCTTTTCGCTTTCTCT 1260
1202 ATAGTGGCGCTACAGAAATTCGAAATTCGCTTTTCGCTTTTCGCTTTCTCT 1260
1261 ATCAGAGAAATTTCTCGAAGAACTCGCTGAGAGACTTGTGATGATGATGATGATG 1320
1262 ATCAGAGAAATTTCTCGAAGAACTCGCTGAGAGACTTGTGATGATGATGATGATG 1320
1321 ATGCGTGGCTCGATGAGTTCGCTTCCGAGGAGTTCACCAAAATATCACTTGGACA 1380
1322 ATGCGTGGCTCGATGAGTTCGCTTCCGAGGAGTTCACCAAAATATCACTTGGACA 1380
1381 AAGAGGT 1387
1382 AACGGTG 1387

```

```

RESULT 8
AAC36324
ID AAC36324 standard; DNA; 1605 BP.
XX AAC36324;
DI 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13376.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.

```

```

OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123254.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0128242.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139484.
PR 18-JUN-1999; 99US-0139485.
PR 18-JUN-1999; 99US-0139486.
PR 18-JUN-1999; 99US-0139487.
PR 18-JUN-1999; 99US-0139488.
PR 18-JUN-1999; 99US-0139489.
PR 18-JUN-1999; 99US-0139490.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139889.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.

```


413 CATCTGAGGCTATAATGCTTGTCTGTTTGGCTTTTCAAAGGAAATGGCAACATAGGAGAA 472
 437 AAGCTGAAGCAACCCGTCGATTAACCAACCAATCTCACCGGAGCCAAATGTTCAASTGT 496
 473 AAGCTCAGGCTACCTATTGATAGCCTAACATTTGCTACTGGGCAATGTTTCAAGTGT 532
 497 GTTGGGAGAAATTCGCTAGTACTTTGAGTGAATTAAGGAAATGAAATTCAGTGAAG 556
 533 GCTGGGAGAAATTTGCAAGTACTTTGAGTAGAGCTCAAAGAGGTGAAATTAAGTGAAG 592
 557 GATACATGATGAGGACCCCTCAACAGCTTTGATATGTTGATGAGAACACCATTTGTTG 616
 593 ACTACTATGTTATGGATCCAGCTTAAAGCTGTAGAGATGGTGGATGAGATACCATCTG 652
 617 TTGGGACATTTCTGTTTCCACTCTTAATCGAGAAATTCGAAGATGTTAAACTCTTGAAC 676
 653 TTGCAGCAATTTAGGATCTACACTTACTGAGAGTTTGGAGCTTTAGCAATTAAGAG 712
 677 ATCTTTGGTCGAAAGAAACAAAGAAACCGGATGGGATACACCAATCCACGTTGATGGG 736
 713 ATCTTTAGCTGAAAGAAACCGGATGGGATACAGGATGGGAACTCTATTTCAATGTTGAT 772
 737 CAAGTGGAGATTCATTTGACCGCTTTTGTATCGGAAATTTGGAATGGGACTTTAGACTTC 796
 773 CCAGTGGAGATTCATTTGCTCCCTTCTTACCTGATCTTGAATGGGACTTTAGGCTTTC 832
 797 CCTTTGGTGAAGATCAATGTTGAGTGGTCCAAAGTATGGACTTTGTTGACGAGGATTTG 856
 833 CATGGTGAAGATTAATACGTCAGTGGTCCAAAGTATGGACTTTGTTGATCGAGGATTTG 892
 857 GTTGGTGAATTCGAGAAACAAAGAGATTTGCTTGAAGAACTATCTTCCATCAATCAAT 916
 893 GTTGGTGTCTGGAGAAACAAAGATTAATTTGCGAGGAACTTTGCTTCCATCAACT 952
 917 ATCTTTGGTGTGACCAACCCACTTTACTCTCAATTTCTCAAAGTTCAAAGTTCAGTCA 976
 953 ACTTGGAGCTGATCAACCCACTTTTCACTTCACTTCAACTTCCAAAGGTTGAGCCAAATCA 1012
 977 TTGCTCAATCACTCAACTTATCCGATTTGGCCACGAGGTTTACAGAAATGATGAGGAA 1036
 1013 TTGCTCAGTACTATCAGTATTCGACTTAGGCTTTGAGGATCAAGAAACATAAATGGAAA 1072
 1097 ATTCAGAGAAATATCGTCTTAAGGAGGACTTTGAGAGACAGAAAGTTTCAACA 1096
 1073 ACTGATGGATACGCAAGGAGGCTAAGAGAAATGAGATGAGAGGAGGAAATTCACA 1132
 1097 TCGTCTCAAAGGACGAGGAGTCCACTTTGCTTTCCTTTCCTTGAAGATGAGGCTGTC 1156
 1133 TTGTTCCAAAGATTTGGCTGCACTTAGTGGATTTCTCTCAAAGACAGTAGCAAGC 1192
 1157 ACATGATTCGAAATCTCCGACATGTTCCGAGGATGATGATGATGATGATGATGATGAT 1216
 1193 ACACGGGTTTGGATGCGAGAGTCTTTGAGAAATTCGGGTGGATCATACCGGCTTACA 1252
 1217 CAATGCTCCAAATGCAACAACATCTCTTTCTGTTGTTTATCAGAGAAATTTCT 1276
 1253 CTATGCTGCAGATGACAGACACATCTGCTGCTCAGAGTTGATGATGATGATGATGAT 1312
 1277 CGAGAACCTCCGTCGAGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1336
 1313 GCCAGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372
 1337 AGCTTCTTCGAGAGTGTATTCACAAATATCACTTTGGACAAGAGAGAGTGAATCTAACA 1396
 1373 GCCTTCTTCGAGAGTGTATTCACAAATATCACTTTGGACAAGAGAGAGTGAATCTAACA 1432
 1397 GCGATACTTGTATG 1410
 1433 AAGTTAAAGTGAAG 1446

AAQ74691 standard; cDNA; 1783 BP.
 AAQ74691;
 25-MAR-2003 (updated)
 22-JUN-1995 (first entry)
 Early Ripening Tomato protein ERTD1 gene.
 Early Ripening Tomato; transgenic plant; crop improvement;
 fruit ripening; transformation; ss.
 Lycopersicon esculentum.
 WO9421794-A1.
 29-SEP-1994.
 22-MAR-1994; 94WO-GB00581.
 22-MAR-1993; 93GB-0005859.
 22-MAR-1993; 93GB-0005860.
 22-MAR-1993; 93GB-0005862.
 22-MAR-1993; 93GB-0005865.
 22-MAR-1993; 93GB-0005866.
 22-MAR-1993; 93GB-0005867.
 22-MAR-1993; 93GB-0005868.
 22-MAR-1993; 93GB-0005869.
 12-JUL-1993; 93GB-0014351.
 12-OCT-1993; 93GB-0020988.
 (ZENE) ZENECA LTD.
 Abu-Bakar UK, Barton SL, Gallego-Veigas PP, Gray JE,
 Grierson D, Lowe AL, Picton S, Whotton LC, Abubakar UK;
 WPI; 1994-317014/39.
 DNA constructs encoding fruit-ripening related proteins - useful
 for transformation of plants to modify fruit quality
 Claim 1; Page 54; 74pp; English.
 ERTD1 homologous transcripts are 1.8 kb in size. The transcript
 is found throughout early fruit development of both wild-type and
 rin (ripening inhibitor) fruit. The transcript is detected
 throughout the entire ripening period in rin fruit. The transcript
 is not detected in leaves or wounded leaves of either rin or wild-
 type tomato. The sequence has been deposited as NCIMB 40588. The DNA
 may be used to transform plants for production of plants with an altered
 level of expression. Improvements in transgenic plants include
 resistance to damage and pests, longer shelf life, improved
 flavor/aroma, etc.
 (Updated on 25-MAR-2003 to correct PN field.)
 Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other;
 Query Match 48.8%; Score 736.6; DB 15; Length 1783;
 Best Local Similarity 71.8%; Pred. No. 2e-215;
 Matches 964; Conservative 0; Mismatches 379; Indels 0; Gaps 0;
 QY 20 TATCGGAGTTCGAGCTTCGGTCCACTCCACATTCGCATCCGTTAGTCCGTTACTTAC 79
 Db 28 TAAGAGATTCAGAGAGAGCTTCGACTGTACTATTGCAUCAAGATATGACAGAACCTT 87
 QY 80 TTCCTAGTTCAGAGATCCGGAACCTCGATTCCTAAGAGAGCGGCTATFCAGATCA 139
 Db 88 TACCTAAGTTCAAAATGCCTAAAAATCCATGCCGAAAGAGAGCTTATCAGATTGTAA 147
 QY 140 ACGAGGAGTATGCTTTCAGCGGATCCAGGTTGAACTTAGCTCTTTGTGACGACAT 199
 Db 148 ACGAGGAGTATGCTTTCAGCGGATCCAGGTTGAACTTAGCTCTTTGTGACGACAT 207

200 GGATGGAGCCCTGAGTGTGATAAATCATCATCATGCTCCATCAACCAAGAACATATGTTGACA 259
 208 GGATGGAGCCCGAGTGGATAGCTCATCATGTCATCCATTAATAAACAATATGTCGACA 267
 260 TGGACGAGTACCCCGTACACCGAATCTCAGAACCGATGCTGTGACATGATGACATFC 319
 268 TGGATGAGTATCTCTGTCACCACTGAACTTCAAAAATAGATGTGTTAACATGTTAGCACATC 327
 320 TATTCAATGACACCGTTAGAAAGGGCGAGACCGCTCGGAGTAGGAACCGTGGATCAT 379
 328 TTTTCCATGCCCGGTGGTGTGATGATGAGCTGCAAGTGGAGTTGGTACAGTGGGTTCA 387
 380 CGGAGGCCATAAATGTTGGCCGTTTGGCCCTTCAAGCGTAAATGGCGAACAAGCGGAAG 439
 388 CAGAGGCAATAAATGCTTGTCTGGCTTCTTCAAGCGCAATGGCAATCGAAAAGAAAG 447
 440 CTGAAGCCAAACCGTGGATAAACCACATGTCACCGGAGCCCATGTTCAAGTGTGT 499
 448 CAGAAAGCAACCTTTCGATAAGCCTAATATAGTCACTGGAGCTAATGTCAGGTTCTGCT 507
 500 GGGAGAAATTCGCTAGCTACTTTGAGGTTGAACTTAAAGGAAGTGAATTTGAGTGAAGAT 559
 508 GGGAAAATTTGCAAGTATTTTGGAGTTGAGTTGAAAGGAGTGAACCTAAAAGAAAGAT 567
 560 ACTATGATGGACCCCTCAACAGCTTGTGATATGTTGATGATGATGATGATGATGATGAT 619
 568 ACTATGATGGACCCCTGCAAAAGCAGTATGATGATGATGATGATGATGATGATGATGAT 627
 620 CGGACATTTTGGTTCACCTTAAATGGAGAACTTGAAGATGTTAACTTCTTGAACGATC 679
 628 CTGCAATCCTTGGTTCCTACTGCTGGAGTTTGGAGTGTGAGCTTCTTAAACGAGC 687
 680 TCTTGGTGGAAAAGAAACAAGAAACCGGATGGGATACCAATFCCACGTTGGATGGCGAA 739
 688 TCCTTACAAAAGAAACAAGAAACCGGATGGGATGGGATGGGATGGGATGGGATGGGAT 747
 740 GTGGAGATTCATTTGACCGCTTTTGTATCCGGAATTTGAAATGGGACTTTAGACTTCCCT 799
 748 GTGGAGATTTATGCTTCTTCTGTCGACATCTTGAATGGGATTTCCGTTTGGCTTC 807
 800 TGGTGAAGATATCAATGTGAGTGGTCAAGATGATGACTTGTACGAGGATGGTGT 859
 808 TTGTGAAAAGTATAAATGTGAGCGGTACCAAGTATGCTTGTATATGCTGTGTGGT 867
 860 GGGTATCTGAGNAACAAGAGGATTTGCTGAGGACTCACTTCTCATATCAATATC 919
 868 GGGTATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927
 920 TTGGTGTGACCAACCCCTTTACTCTCAATTTCTCAAAGGTTCAAGTCAAGTCAATG 979
 928 TTGGTGTGATCAGCTTACTTCTCAACTTCTTAAAGGTTTCTTAAATTAATTTG 987
 980 CTCNAATCACTCAACTTATCCGATGGGCCACAGGGTTACAGAAATGATGAGGAAAT 1039
 988 CACAGTATTAATCAAGTAAAGACTTGGCTTTTGAAGGTTTATAAGAACTGATGAAGAAT 1047
 1040 GCAGAGAAATATCATGCTCTAAGGAAGACTTGAAGAAAGCAAGAAAGTTCAACATCG 1099
 1048 CTTTATCAACCGCAAAAGTACTACAGAGGAAATCAAAAATGGGGCGTTCGATTTG 1107
 1100 TCTCAAAGGACGAGGAGTGCCTTCTGCTTCTTCTTAAAGATAGCAGCTGTCA 1159
 1108 TCTCAAAGGATGGGTGTCTCTGTTAGCAATTTCTCAGGGACAGCAGCAATAATA 1167
 1160 CTGAGTTCGAAATCTCCGACATGTTGCGAGGATGATGATGATGATGATGATGATGATGAT 1219
 1168 CGGATTTGAAGTATCTGAGCAATCTGAGAAGATTTGATGATGATGATGATGATGATGAT 1227
 1220 TCCCTCCAAATGCAACAACATCACTTCTTCTGTTGTTTATCAGAGAAATTTCTCGA 1279
 1228 TGCCACCGGATGCTGAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1287
 1280 GAACACTGCTGAGAGACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1339

1288 ACAGCCTAGCTGAGAGACTTGTTCGACATTTGAGAAAATTTCTGTGAGATGGACACAC 1347
 1340 TTCCTTGGAGATGATTCACAAA 1362
 1348 AGCCTCTGTTTGGCCACAAA 1370

RESULT 10
 AAL43418
 ID AAL43418 standard; cDNA; 1783 BP.
 XX AC AAL43418;
 XX DT 25-SEP-2002 (first entry)
 XX Tomato GAD coding sequence.
 DE GAD; plant GABA production regulation; glutamic acid decarboxylase;
 KW plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
 XX Lycopersicon esculentum.
 EH Key Location/Qualifiers
 FT CDS 6..1514
 FT /*tag= a
 FT /product= "GAD"
 XX WO200238736-A2.
 XX 16-MAY-2002.
 XX 07-NOV-2001; 2001WO-US47447.
 XX 07-NOV-2000; 2000US-246367P.
 XX (EMER-) EMERALD BIOAGRICULTURE CORP.
 XX Kinnersley AM, Turano FU;
 XX WPI; 2002-490073/52.
 XX P-PSDB; AAO15140.
 XX Making transgenic plants that selectively increase gamma-aminobutyric
 acid production, by incorporating a DNA construct with a polynucleotide
 encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 59; 63pp; English.
 The present invention relates to a method of producing a transformed
 plant that selectively increases production of gamma-aminobutyric acid
 (GABA) in response to a signal, by incorporating into the plant's genome
 a DNA construct with a non-constitutive promoter operably linked to a
 polynucleotide encoding a functional plant glutamic acid decarboxylase
 (GAD), to provide a transformed plant that expresses the GAD coding
 sequence in response to a signal. Plants of this type have an enhanced
 ability to tolerate environmental or other stresses. The present sequence
 is the tomato GAD coding sequence.

Query Match 48.8%; Score 736.6; DB 24; Length 1783;
 Best Local Similarity 71.8%; Pred. No. 2e-215;
 Matches 964; Conservative 0; Mismatches 379; Indels 0; Gaps 0;
 20 TATCGGAGTGGGACGTTCCGTTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 79
 28 TAAGAGATTCAGAAGAGAGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 87
 80 TTCCTAGTTCAGATGTCGCGAAAATTCGATTCCTTAAGGAAGCGGCTATCAGATCATCA 139
 88 TACCTAAGTTCAAAATGCTTAAAAAATCCATGCGCAAGACGCTTATCAGATTTAA 147

140 ACACGAGCTGATGCTTTCGCGGAAATCCACGGTTGAACTTAGCCTCCCTTTGTCAGCAGAT 199
 148 ACACGAGCTTATGTTGGATGTAACCCAGGTTGAAATTTAGCTTCTTTGTTAGCAGAT 207
 200 GGATGAGCCCTGATGATGATAAATCAATCATGTCCTCCATCAACAGAACTATGTTGACA 259
 208 GGATGGAGCCCGAGTGGATAAGCTCATCATGTCATCCATTAATAAATAAATACTATGTCGACA 267
 260 TGGACGAGTACCCCGCTCACACCGAATTCAGAAACCGATGTTGAACTATGTTGACATC 319
 268 TGGATGAGTATCTCTGACCACTGAACTTCAAAATAGATGTTAACTAATGTTAGCAGATC 327
 320 TATTCAATGACCCGTTAGAGAGCGGAGACCGCCGCTCGGAGTAGAACCCGTTGGATCAT 379
 328 TTTTCCATGCCCGGTTGGTGTGATGATGAGACTCAGTTGGAGTTGGTACAGTGGGTTTCA 387
 380 CGGAGCCATAAATGTTGGCCGTTTGGCCCTTCAAGCGTAAATGGCAGAACAGCCGAAG 439
 388 CAGAGGCAATAATGCTTGTCTGGCCCTTGCCTTTCAAAACGCAAAATGGCAATCGAAAGAAAG 447
 440 CTGAAGCCAAACCCGTCGATFAAAACCCAAATTTGTCACCGGAGCCAAATGTTCAAGTGTGTT 499
 448 CAGAGGCAAACTTTCGATGAGCCCTAATATATGATCACTGGAGCTAATGTCAGGTTCTGCT 507
 500 GGGAGAAATTCGCTAGTACTTTGAGGTTGAACTTAAAGAAATGAAATTTGAGTGAAGAT 559
 508 GGGAAAAATTTGCAAGGTAATTTTGGGTTGAGTTGAAAGGAGTGAATACTAAAGAAAGAT 567
 560 ACTATGTTGAGGACCTCAACAAGCTTGTATATGTTGATGAGAACACCAATTTGTTGTTG 619
 568 ACTATGTAATGGACCTTCCAAAGCAGTAGAGATGTTGGATGAGAAATCAATATGTTGTTG 627
 620 CGGACATCTTGGTTCCACTTCTAATGGAGAAATCGAAGATGTTAAACTCTTGAACGATC 679
 628 CTGCAATCTTGGTTCTACTCTGACTGGGAGTTTGAGGATGTTGAAAGCTCTAAACGAGC 687
 680 TCTTGTGCAAAAGAACAAAGAAACCGGATGGATACCAATCCACGTTGGATCGCGGAA 739
 688 TCTTACAAAAAAGAACAAAGAAACCGGATGGGAGACCCGATTCATGTCGATGTCGGA 747
 740 GTGGAGGATTCATGACCCGCTTTTGTATCCGGAATTTGAAATGGGACTTTAGACTTCCCT 799
 748 GTGGAGGATTTATGCTTCTTCTGCGCAGATCTTGAUGGAAATTTCCGTTCCCTC 807
 800 TGGTGAAGATCAATGATGAGTGTGTCACAAAGTATGGACTTTGTTGTAACGAGGATGGTT 859
 808 TTGTGAAAAGTATAAATGTCAGCGGTACCAAGTATGGCTTGTATATGCTGTGTTGCGTT 867
 860 GGGTATCTGGAGAAACAAAGAGATTTGCTGAGGAACTCATCTTCCATATCAATATC 919
 868 GGGTATGATGGCGGAAAGAGACTTTGCCGATGAACTTGGCTTTTTCATATAAACTACC 927
 920 TTGGTGTGACCAACCCACCTTTACTCTCAATTTCTCAAAAGGTTCAAGTCAAGTCAATG 979
 928 TTGGGCTGATCAGCCTACTTTTACTCTCACTTCTTAAAGTTCTTCAATCAATATG 987
 980 CTCAATCTACCAACTTATCCGATGCGGCAAGGTTTACAGAAATGATGATGGAGAAAT 1039
 988 CACAGTATATCAGTTAATAAGACTTGGCTTTTGGGTTTATAAAGACGTCATGAAGAAAT 1047
 1040 GCAGAGAGATATGATCGTCTTAAGGAGAGACTTTGAGAGAGACAGAAAGTTCAACATCG 1099
 1048 CTTATCAACCGCAAAAGTACTCAACAGAGGAAATCAAAAATATGGGCGGTTTGGATTTG 1107
 1100 TCTCAAAAGGACGAGGATGCGCACTTCTGCTTTCTCTTGAAGATGAGCAGCTGTCA 1159
 1108 TCTCTAAGGATGTTGGTGTCTTGTAGCAATTTCTCTCAGGACAGCAGCAATATA 1167
 1160 CTGAGTTGCAAAATCTCCGACATGCTTCCGAGGATGATGGAATGATGTCGCGGCTTACAAA 1219
 1168 CCGTATTTGAGATCTGAGCATCTCAGAAAGATTTGGATGGATCGTCCCTGTCATACAAA 1227

1220 TGCCTCCAAAATGCAACAACACATCACTGTTCTTCGTTGGTATTACAGAAAGATTTCTCGA 1279
 1228 TGCACCCGGATGCTGAACACACATTCCTACTACTGCGGTTTGTCAATTAGAGAGGATTTACGCC 1287
 1280 GAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTATGCGCTGAGCTCGATGAGC 1339
 1288 ACAGCTAGCTGAGAGACTTGTTCAGATTTGACATTTGAGAAAATTTCTGTCAGATTTGGACAC 1347
 1340 TTCCTTCGAGAGTGAATTCACAAA 1362
 1348 AGCCTCCTCGTTGCCACCAAAA 1370

RESULT 11
 AAI70004
 ID AAI70004 standard; DNA; 1180 BP.
 XX AAI70004;
 AC AAI70004;
 XX
 DT 24-JAN-2002 (first entry)
 DE Tomato glutamate decarboxylase clone T-gad-19 DNA sequence.
 KW Tomato; glutamate decarboxylase; GAD; transgenic plant; T-gad-19; ds.
 XX
 OS Lycopersicon esculentum.
 XX
 FN JP2001238555-A.
 XX
 PD 04-SEP-2001.
 XX
 PF 13-DEC-2000; 2000JP-0404172.
 XX
 PR 13-DEC-1999; 99JP-0352552.
 XX
 PA (AJIN) AJINOWOTO KK.
 XX
 DR WPI; 2001-610050/70.
 XX
 PT Creation of a transgenic plant of improved amino acid composition, involves transformation with antisense sequences -
 XX
 PS Example 1; Page 13-14; 18pp; Japanese.
 XX
 CC The present invention relates to a method for creating a transgenic plant which accumulates a larger amount of at least one amino acid of edible part than a natural plant of the same species cultured under a same condition. The method involves transforming a plant with a genetic structure containing an antisense sequence of a gene sequence encoding glutamate decarboxylase (GAD) and a controlling sequence which can express the antisense sequence. The resultant transgenic plant has high added value. The present sequence is the DNA sequence for clone T-gad-19 of the tomato GAD gene, which was used to illustrate the present invention.
 XX
 SQ Sequence 1180 BP; 372 A; 188 C; 276 G; 344 T; 0 other;

Query Match 47.1%; Score 711.4; DB 22; Length 1180;
 Best Local Similarity 75.7%; Pred. No. 9.1e-208;
 Matches 894; Conservative 0; Mismatches 286; Indels 1; Gaps 1;

QY 11.2 CCTAAGGAGCGGCTATCAGATCATCAACGACGCTGATGCTTGAACGGAATCCACGG 171
 Db 1 CCAAGGAGGAGCAGATCAGATTTCTTAATGATGACTTATGTTAGATGGAATCCAGG 60
 QY 172 TTGAATTTAGCCTCTTTGTGACGACATGATGGAGCTGAGTGTGATAAATCATCATG 231
 Db 61 TTGAATTTGGCATCTTTTGTCAACATGATGAAACGAGAAATGCAAAATTTGATGATG 120
 QY 232 TCCCTCATCAACAGAACTATGTCAGTGGACGAGTACCCCTCACCACCGAACTTCAG 291
 Db 121 GATTCATTAACAAAAATTTATGTTGACATGGATGAATATCTGTCACCACCTGAGCTTCAG 180

```

292 AACCGATGTTGAACAGTAAAGTTCACATATCAATGCACCCGTTAGAGAGCGGAGACC 351
181 AATCGGTGTGTAACATGATAGCGCATTTAATTAATGCACCACTTGAAGATGGAAACT 240
352 GCGTTCGGAGTAGGACCCCTGATCATCGGAGCCATATGTTGCCGGTTGGCCCTC 411
241 GCAGTTGGATTGGAAAGTGGTTCTTTCAGAAAGCCATTATGTTGCTGGATTGGCC 300
412 AACCGTAAATGGCAGAACCAAGCCCAAGAGTGAAGCCAAACCCGTCGATAAACCCCA 471
301 AGAGAAATGGCAACCAAAATGAAGCCCAAGAGCCCTATGATTAAGCCCAACT 360
472 GTCACCGGAGCCAAATGTTCAAGTGTGTGGGAAATTCGCTAGGTAATTTGAGGTTG 531
361 GTTACTGGTCTAATCTCCAGGTGTGTGGGAAATTTGCAAGGTATTTTGAAGTTG 420
532 CTTAAGGAGTGAATTTGAGTGAAGATCATATGATGACCCCTCAACAAGCTGTTGAT 591
421 CTRAAAAGAGTGAAGTTGACTGATGATCATATGATGAGCCCTGAGAAAGCTG 480
592 ATGGTTGATGAGAACCACTTTGTTGGGCAATTCCTGGTTCACCTCTTAATGGAGAA 651
481 ATGGTTGATGAGAACCACTTTGTTGGGCAATTCCTGGTTCACCTCTTAATGGAGAA 540
652 TTCGAAAGTGTAACTCTTGAAGATCTTGGTGTGAAAGAAAGAAAGAAAGCCGATGG 711
541 TTTGAGGAGTGAAGAAATTTGAATGACCTCTCT-ATTGAAAGAAAGAAAGAAAGCCG 599
712 GATACACCAATCCAGTGGATGGGCAAGTGGAGGATTCATTCACCGTTTTGATCCG 771
600 GACACTCAATTCATGTTGATGAGCTAGTGGTGGATTTATGACCACTTATATATCA 659
772 GAATGGAATGGGACTTTAGACTTCCCTGGTGAAGAGATCAATCAAGTGTGATGAT 831
660 GAACCTTGAATGGGACTTTAGACTTCCCTGGTGAAGAGATCAATCAAGTGTGATG 719
832 TATGGATGTTGATGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 891
720 TATGGCCTTGTATATGTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 779
892 GAGGAATCATCTCCATATCAATTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 951
780 GATGAACTTATTTTCAATTAATTTACCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 839
952 TTTCCAAAGGTTCAAGTCAAGTCAATGCTCAATGCTCAATGCTCAATGCTCAATG 1011
840 TTTCTAAAGGATCAAGTCAAGTCAATGCTCAATGCTCAATGCTCAATGCTCAATG 899
1012 GAGGTTACAGAAATGATGAGGAAATGCGAGAGAAATGATGATGATGATGATGATG 1071
900 GAGGTTAAGAAACGTCATGAAGAAATGCTTATCAACCAAGGTTACTCAAGAGGA 959
1072 CTTGAGAGACAGAAAGTTCAACATGCTCTCAAGGACGAGGAGTGGCCACTGTCGCT 1131
960 ATCAAAAATGGGAAGATTCGAGATGTTCTCAAGAGTGTGGTGTTCCTTGGTTGCA 1019
1132 TTTCTTGAAGATGACGTTGACACTGAGTTGAAATTCGCAATGCTCCGACATGCTCCG 1191
1020 TTTTCTTAAAGCAATAGCAAAATGATGATGATGATGATGATGATGATGATGATG 1079
1192 TATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
1080 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
1252 CGTGTGTTATCAGAGAGATTTCTCGAGAACACTCGCTGA 1292
1140 AGAGTTGTTATTAGAGAGATTTCTCCCGTACACTCGCCGA 1180

```

```

AC ABX62354;
XX 25-FEB-2003 (first entry)
XX Arabidopsis thaliana expressed sequence related polynucleotide #469.
DE Transgenic plant; plant; genetically modified cell;
XX environmental stress; ribozyme creation; disease resistance;
KW stress tolerance; fungicide screening; insecticide screening;
KW gene; ds.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX US2002040490-A1.
XX 04-APR-2002.
XX 26-JAN-2001; 2001US-0770423.
XX 27-JAN-2000; 2000US-178512P.
XX (GORL/) GORLACH J.
XX (ANVY/) AN Y.
XX (HAM/) HAMILTON C M.
XX (PRIC/) PRICE J L.
XX (RAIN/) RAINES T M.
XX (YUYU/) YU Y.
XX (RAME/) RAMEAKA J G.
XX (PAGE/) PAGE A.
XX (MATH/) MATH A V.
XX (LEDF/) LEDFORD B L.
XX (WOES/) WOESSNER J P.
XX (HAAS/) HAAS W D.
XX (GARC/) GARCIA C A.
XX (KRIC/) KRICKER M.
XX (SLAT/) SLATER T.
XX (DAVI/) DAVIS K R.
XX (ALLE/) ALLEN K.
XX (HOFF/) HOFFMAN N.
XX (HURB/) HURBAN P.
XX Goriach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
XX Ramaska JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
XX Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
XX Hurban P;
XX WPI; 2003-110411/10.
XX Novel Arabidopsis thaliana nucleic acid useful for identifying
XX homologous or related genes, and to create genetically modified and
XX transgenic organisms, such as plant cells and plants -
XX Claim 1; SEQ ID NO 469; 43pp; English.
XX The invention describes an Arabidopsis thaliana nucleic acid (I). The
XX polypeptide (II) encoded by (I), transgenic plant (III) or genetically
XX modified cell (IV) are useful for screening a candidate agent for its
XX biological effect, by combining the candidate agent with (II), (III) or
XX (IV), and determining the effect of the candidate agent on (II), (III) or
XX (IV). (I) is useful for identifying homologous or related genes, for
XX producing compositions that modulate the expression or function of its
XX encoded protein, for mapping functional regions of the protein, in
XX diagnosis, for studying associated physiological pathways, for genetic
XX manipulation of cells, preferably plant cells, in screening assays of
XX various plant strains to determine the strains that are capable of
XX withstanding a particular disease or environmental stress, for enhancing
XX or inhibiting production of biosynthetic product in a plant, for
XX producing polypeptides as probes for the detection of mRNA in biological
XX samples, to generate additional copies of (I), to generate ribozymes or
XX oligonucleotides, as single stranded DNA probes or as triple-strand
XX forming oligonucleotides, and to create genetically modified and
XX transgenic organisms, such as plant cells and plants. (II) or (III) is
XX useful for introducing or improving disease resistance and stress

```



```

16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
18-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145066.
22-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
02-AUG-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157153.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.

```

```

PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 25.88; Score 388.6; DB 21; Length 1607;
Best Local Similarity 64.9%; Pred. No. 1.9e-108;
Matches 793; Conservative 0; Mismatches 284; Indels 144; Gaps 9;

QY 292 AACCGATGTGTGAACATGATGACATCTATTCATATGCACCGTTAGAAGGGGGAGACC 351
DB 361 AACCGATGTGAACATTAATAGCTCGACTTTCATATGCGCCACTCGAGGAATCGAGAG 420
QY 352 GCCGTGGAGTAGAAACCGTTGGATCATCGAGGCCATAATGTTGGCCGGTTTGGCCCTTC 411
DB 421 CGCGTGGAGTAGGACAGTTGGTCTTCAGAAAGCCATCATGTTAGCCGGATTTGGCCCTTC 480
QY 412 AAGCGTAAATGGCAGAACAGCCCAAGCTGAAGGCAAAACCCGTCGATAAACCCAACTTT 471
DB 481 AAAGAAAATGGCAGAACAAACCGAAGCTGAGGTAACCCCTATGCAAAACCCAACTTT 540
QY 472 GTCACCGGAGCCAAATGTTCAAGTGTGTTGGGAAATTCGCTAGGTACTTTGAGTTGAA 531
DB 541 TGTCACCGGATGCCAATCAAGTTTGGTTGGAGAAATTCGCTGGTACTTTGAGGTGGAG 600
QY 532 CTTAAGGAAGTGAATGAGTGAAGGATCTATGATGGACCCCTCAACAGCTGTTGAT 591
DB 601 C---AAGGATTAACCTAAGTGAAGTACTACTGATGGATCCAGCAAAAGCAGAGAA 657
QY 592 ATGGTTGATGAGAACACCAATTTGTTGGGCAATTTGTTGGTCCACTCTTAATGAGAA 651
DB 658 ATGGTAGCAGAACCAATCTGTTCGCA--ATATGGGATCCACTCAACGCA--- 711
QY 652 TTCGAGATGTTAACTCTTGTGAACGATCTTGTGTCGAAAAGAAACAAAGAACCCGGATGG 711
DB 712 ----- 711
QY 712 GATACACCAATCCACGTTGGATGGCCAGTGGAGGATTCATTCACCGCTTTTGTATCCG 771
DB 712 -----AGTGGGTATAGCTCCGCTTTATCTCT 741
QY 772 GAATTGGAATGGACTTTAGACTTCCCTTGGTGAAGATCAATGATGAGTGTCAAG 831
DB 742 GAATTGGAATGGACTTTAGCTTCCCTTGGTGAAGATCAACCTGAGTGTGTCACAG 801
QY 832 TATGGACTTGTACCGCAGGATGTTGTTGGTGTATCTGGAGAAACAAAGGATTTGCTT 891
DB -----

```


Y 889 CCGAGAACTCTTCCATATCAATATCTGGTGTGACCAACCCACTTTACTTTC 948
 |||||
 b 940 CCCAAGGATCGATTTTCAACTCACTACCTAGGAGCTGAACAGCCAGTTTACCCCTG 999
 |||||
 Y 949 AATTTCTCCAAAGGTTCAAGTCAAGTCAATGTCACACTTATCCGATTTGGCC 1008
 |||||
 b 1000 AACTTCTCCAAAGGTTCAAGTCAAGTCAATGTCACACTTATCCGATTTGGCC 1059
 |||||
 Y 1009 CACGAGGTTTACAGAAATGATGGAGAAATGACAGAGAAATATGATCGTCTTAAGGGAA 1068
 |||||
 b 1060 AAGCGTGGCTACCGCTCCATCATGACCAACATATGTTACTGGGACTTTTGGCCCAA 1119
 |||||
 Y 1069 GGACTTGGAGA-----CAGAAAGGTTCAACATGCTCTCAAAGGACGAGGAGT 1118
 |||||
 b 1120 GAGCTGGAGAGATGGCTTTATCATATGAGCCAGCTGCGCGCCACGGTCTCCCAATTG 1179
 |||||
 Y 1119 GCCACTTGTGGCTTTCTCTTGAAGATACAGCTGTCACACTGAGTTCGAAATCTCCGA 1178
 |||||
 b 1180 GTGCTTTGACTGCCCCGTGAGCGGAGGGTCAAGTTCGAC--GAGTTTGCCTTTGGCACA 1238
 |||||
 Y 1179 CATGCTTCGAGGTAAGGATGATGAGTCCCGGCTACACAAATGCTCCAAATGCACAACA 1238
 |||||
 b 1239 CCAGCTCGTGGCTGGATGTCCTGTCATACAGATGGGCCCCAACAGCAACA 1298
 |||||
 Y 1239 CATCACTGTTCTTGGTGTATCAGAGAGATTTCTCGAAGACTCGCTGAGAGACT 1298
 |||||
 b 1299 CTTGAAGCTGATCGGTGTTGTCGCGAGGATTCACAAAGAGCGGTGCGAGCGGCT 1358
 |||||
 Y 1299 TGTGATCGATAGAGAAAGTATGCTGAGCTCGATGA 1337
 |||||
 b 1359 GCTTAGTATCAAGTTGGCTTGAAGACTTGGGTGA 1397
 |||||

CC (GABA) in response to a signal, by incorporating into the plant's genome
 CC a DNA construct with a non-constitutive promoter operably linked to a
 CC polynucleotide encoding a functional plant glutamic acid decarboxylase
 CC (GAD), to provide a transformed plant that expresses the GAD coding
 CC sequence in response to a signal. Plants of this type have an enhanced
 CC ability to tolerate environmental or other stresses. The present sequence
 CC is the A. thaliana GAD4 coding sequence.
 XX
 SQ Sequence 2121 BP; 626 A; 412 C; 473 G; 610 T; 0 other;
 Query Match 25.6%; Score 385.8; DB 24; Length 2121;
 Best Local Similarity 65.6%; Pred No. 1.6e-107;
 Matches 701; Conservative 0; Mismatches 182; Indels 186; Gaps 2;
 QY 84 TAGGTTCAAGATGCCGGAATACTCGATTCTTAAGGAAGCGGCTATCAGATCATCAAGA 143
 Db 270 TAGATTGGAATGCCGGAATACTCAATCCAAAAGAAGCAGCTTACCAAAATCATCAAGA 329
 QY 144 CGAGCTGATGCTTGAAGGAAATCCAGGTTGAACTTAGCTTCCCTTTGACGACATGAT 203
 Db 330 CGAGCTAATGCTTCGATGTAACCCAAAGCTGAACCTAGCTTCCCTCGTACCACATGAT 389
 QY 204 GGAGCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAGAACTATGTTGACATGA 263
 Db 390 GGAGCCAGAAATGTGACCAACTCATGATGAGTCCATCAACAGAACTAGCTCGACATGA 449
 QY 264 CGAGTACCCCGTCAACCCGAACTTCGAAACCGATGTTGTAACATGATTGCACTACTATT 323
 Db 450 CGAGTACCCCGTCAACCCGAACTTCGAAACCGATGTTGTAACATGATTGCACTACTATT 509
 QY 324 CAATGCAACGTTAGAGAGGCGGAGACCCCGCTCGGAGTAGGAACCGTTGATCATCGGA 383
 Db 510 CAACCGCGCGTTCGATGACCGGTGAAGCTGCGCTTGGTGTGGCACCCGTCGGATCGGA 569
 QY 384 GGCCATAATGTTGGCGGTTTGGCTTCAAGCGTAAATGGCGAACAAGCGCAAGCTGA 443
 Db 570 GGCGATATGTTGGCGGTTTGGCTTAAAGACAATGGCAGAAATGAGCGTAAAGGCCCA 629
 QY 444 AGGCAAAACCGTFCGATAAAACCCCAACTTGTCAACCGGACCAATG-----487
 Db 630 AGGGTTCCTTATGATAAGCCCAATATCGTAAACCGGTGCTAATGTCCAGGTAACCAAAA 689
 QY 488 -----487
 Db 690 CAAAATGATGAAATATTAACCAAGACAAAATGAAATTAATCAATCCCGTTAAGTTATA 749
 QY 488 -----TTCAAGTGTGTGGGAGAAAATCGTAGTACTTTG 523
 Db 750 TGTGTGACTCAATTTCCGGTTCATACAGTTTGTGGAGAAATTCGAAAGTATTTCCG 809
 QY 524 AGGTTGAACTTAAGGAAGTAAATGAGTGAAGGATACTATGTGATGAGACCTCAACAG 583
 Db 810 AAGTGGAGCTTAAGGAAGTAAACCTTAAGAGAAAGACTATTAAGTATGAGACCTGTAAGG 869
 QY 584 CTGTTGATATGTTGATGAGAACACTTGTGTTGCGGACATTTCTTGGTCCACTCTTA 643
 Db 870 CGGTGAAATGTTAGACGAAAACAAATTTGTGTCGCTGCCATCTCGGTTCAACGTTAA 929
 QY 644 ATGAGAAATTCGAGATGTTAAACTTCTTGAACGATCTTGTGCGAAAAGAACAAAGAAA 703
 Db 930 CGGTTGAAATTCGAGACGTTAAGCTCCTCAACGACCTCTTGTGCGAGAAAACAAGCAA 989
 QY 704 CCG-----706
 Db 990 CCGGTAATTAACCAACCGAACAAGCTAATCGATTGTAATCCGTTTGGGTTGGAGTCC 1049
 QY 707 -----GATGGATACACCAATCCACCGTGGATCGGC 737
 Db 1050 GGTTTTAAAGCTTAAACACAAATTTGCGATGGGACACCCCAATACACGTTGGACGACG 1109
 QY 738 AAGTGGAGGATTCATTGACCGGTTTTTGTATCCGGAATGGAATGGGACTTTAGACTTCC 797
 Db 1110 GAGTGGTGGTTTATGTTCCGTTCTTGTATCCGAGGCTGGAGTGGGATTTCCGGCTACC 1169

RESULT 15
 AAL43413
 AAL43413 standard; DNA; 2121 BP.
 AAL43413;
 25-SEP-2002 (first entry)
 A thaliana GAD4 coding sequence.
 GAD; plant GABA production regulation; glutamic acid decarboxylase;
 plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
 Arabidopsis thaliana.
 W0200238736-A2.
 16-MAY-2002.
 07-NOV-2001; 2001WO-US47447.
 07-NOV-2000; 2000US-246367P.
 (EMER-) EMERALD BIOAGRICULTURE CORP.
 Kinnersley AM, Turano FJ;
 WPI: 2002-490073/52.
 P-PSDB; AAO15135.
 Making transformed plants that selectively increase gamma-aminobutyric
 acid production, by incorporating a DNA construct with a polynucleotide
 encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 55; 63pp; English.
 The present invention relates to a method of producing a transformed
 plant that selectively increases production of gamma-aminobutyric acid

```

y 798 CTTGGTGAAGAGTATCAATGTGAGTGGTCAAAAGTATGGACTTGTGTACCCAGGGATTGG 857
o 1170 GTTGGTTAAGAGTATTAAATGTGAGTGGTCAAAATACGGTTTGGTTTACGCCGGTATTGG 1229
y 858 TTGGGTGATCTGGAGAAACAAAGAGGATTTGCCCTGAGGAACCTCATCTTCCATATCAATTA 917
b 1230 TTGGGTTGTATGGAGAACCAAAACCCGATTTGCCCTGATGAACCTTATCTTCCATATCAATTA 1289
y 918 TCTTGGTGTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCA 966
b 1290 TCTTGGGCTGATCAACCAACCTTTACTCAACTCACTTCTCCAAAGGTACA 1338

```

earch completed: October 22, 2003, 11:12:29
 ob time : 444 secs

GenCore version 5.1.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

4 protein - protein search, using sw model

in on: October 22, 2003, 13:37:03 ; Search time 38 Seconds
 (without alignments)
 621.248 Million cell updates/sec

File: US-10-006-852-2
 Effect score: 2615
 Sequence: 1 MYLSHAVSESDVSHSTFAS.....DIITGKKFVADRKKTSGIC 502

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database: SwissProt_41.*

Result No.	Score	Query Match %	Length	DB ID	Description
1	2607	99.7	502	DCE1_ARATH	Q42521 arabidopsis
2	2250	86.0	500	DCE_PETHY	Q07346 petunia hybr
3	2130	81.5	494	DCEP_ARATH	Q42472 arabidopsis
4	2014	77.0	502	DCE_LYCES	P54767 lycopersico
5	1028	39.3	466	DCE_LACIA	Q9C920 lactococcus
6	1023	39.1	466	DCE_LACLC	O30418 lactococcus
7	1013	38.7	466	DCEA_ECOLI	P80063 escherichia
8	1013	38.7	466	DCEB_ECOLI	P28302 escherichia
9	1009	38.6	466	DCEA_ECO57	P58228 escherichia
10	963	36.8	467	DCEC_LISIN	Q928K4 listeria in
11	955.5	36.5	462	DCEA_LISMO	O915P3 listeria mo
12	953.5	36.5	467	DCEC_LISMO	O8Y4K4 listeria mo
13	948	36.3	464	DCEB_LISIN	O928R9 listeria in
14	940	35.9	464	DCEB_LISMO	O9eyw9 listeria in
15	887.5	33.9	585	DCE_YEAST	O04792 saccharomyc
16	225	8.6	396	OY050_METJA	O60358 methanococ
17	137	5.2	413	DCHS_LYCES	P54772 lycopersico
18	127.5	4.9	500	DDC_CATRO	P17770 catharanthu
19	124.5	4.8	386	DCHS_VTBAN	O56581 vibrio angu
20	124.5	4.8	510	DDC_ACIBA	O43908 acetobact
21	123	4.7	480	DDC_HUMAN	P20711 homo sapien
22	123	4.7	1935	MYH7_HUMAN	P12883 homo sapien
23	123	4.7	1935	MYH7_PPG	P79293 sus scrofa
24	121.5	4.6	511	DDC_HAEIN	P71362 haemophilus
25	119	4.6	510	DDC_DROME	P18486 drosophila
26	115.5	4.4	656	DCHS_RAT	P16453 rattus norv
27	114.5	4.4	756	METE_BUCBP	O89b24 buchnera ap
28	113.5	4.3	804	YNA4_CAEEL	P45895 caenorhabdi
29	113	4.3	1934	MYH7_MESAU	P13540 mesocriceti
30	112	4.3	593	DDP1_PEA	P51850 pisum sativ
31	112	4.3	1935	MYH7_RAT	P02564 rattus norv
32	111.5	4.3	377	DCHS_KLEPL	P28578 klebsiella
33	110.5	4.2	1420	APX_XENLA	Q01613 xenopus lae

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2607	99.7	502	DCE1_ARATH	Q42521 arabidopsis
2	2250	86.0	500	DCE_PETHY	Q07346 petunia hybr
3	2130	81.5	494	DCEP_ARATH	Q42472 arabidopsis
4	2014	77.0	502	DCE_LYCES	P54767 lycopersico
5	1028	39.3	466	DCE_LACIA	Q9C920 lactococcus
6	1023	39.1	466	DCE_LACLC	O30418 lactococcus
7	1013	38.7	466	DCEA_ECOLI	P80063 escherichia
8	1013	38.7	466	DCEB_ECOLI	P28302 escherichia
9	1009	38.6	466	DCEA_ECO57	P58228 escherichia
10	963	36.8	467	DCEC_LISIN	Q928K4 listeria in
11	955.5	36.5	462	DCEA_LISMO	O915P3 listeria mo
12	953.5	36.5	467	DCEC_LISMO	O8Y4K4 listeria mo
13	948	36.3	464	DCEB_LISIN	O928R9 listeria in
14	940	35.9	464	DCEB_LISMO	O9eyw9 listeria in
15	887.5	33.9	585	DCE_YEAST	O04792 saccharomyc
16	225	8.6	396	OY050_METJA	O60358 methanococ
17	137	5.2	413	DCHS_LYCES	P54772 lycopersico
18	127.5	4.9	500	DDC_CATRO	P17770 catharanthu
19	124.5	4.8	386	DCHS_VTBAN	O56581 vibrio angu
20	124.5	4.8	510	DDC_ACIBA	O43908 acetobact
21	123	4.7	480	DDC_HUMAN	P20711 homo sapien
22	123	4.7	1935	MYH7_HUMAN	P12883 homo sapien
23	123	4.7	1935	MYH7_PPG	P79293 sus scrofa
24	121.5	4.6	511	DDC_HAEIN	P71362 haemophilus
25	119	4.6	510	DDC_DROME	P18486 drosophila
26	115.5	4.4	656	DCHS_RAT	P16453 rattus norv
27	114.5	4.4	756	METE_BUCBP	O89b24 buchnera ap
28	113.5	4.3	804	YNA4_CAEEL	P45895 caenorhabdi
29	113	4.3	1934	MYH7_MESAU	P13540 mesocriceti
30	112	4.3	593	DDP1_PEA	P51850 pisum sativ
31	112	4.3	1935	MYH7_RAT	P02564 rattus norv
32	111.5	4.3	377	DCHS_KLEPL	P28578 klebsiella
33	110.5	4.2	1420	APX_XENLA	Q01613 xenopus lae

ALIGNMENTS

RESULT 1

ID	DCE1_ARATH	STANDARD	PRT	502 AA
AC	Q42521	Q9FFH9		
DT	01-NOV-1997	(Rel. 35, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Glutamate decarboxylase I (EC 4.1.1.15) (GAD I).			
GN	GADI OR GDH1 OR GAD OR AT5G17330 OR MKP11.30 OR MKP11_18.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI_taxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=95334488; PubMed=7610159;			
RA	Araki T., Baum G., Snedden W.A., Sheld B.J., Fromm H.;			
RT	"Molecular and biochemical analysis of calmodulin interactions with the calmodulin-binding domain of plant glutamate decarboxylase.";			
RL	Plant Physiol. 108:551-561(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=97471969; PubMed=9330910;			
RA	Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M., Miyajima N., Tabata S.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned P1 clones.";			
RL	DNA Res. 4:215-230(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;			
RT	"RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the SSP consortium (Salk/Stanford/PGEC)";			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
CC	!- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS WAY, DIRECTLY OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA BIOSYNTHESIS.			
CC	!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).			
CC	!- COFACTOR: Pyridoxal phosphate.			
CC	!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC, GAD, HDC AND TRDCC).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			

```

RT "A plant glutamate decarboxylase containing a calmodulin binding
RT domain. Cloning, sequence, and functional analysis."
RL J. Biol. Chem. 268:19610-19617 (1993).
CC !- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING
CC IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS MAY, DIRECTLY
CC OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA
CC BIOSYNTHESIS.
CC !- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC !- COFACTOR: Pyridoxal phosphate.
CC !- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
CC GAD, HDC AND TYRDC).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L16797; AAA33709.1; -
DR EMBL; L16977; AAA33710.1; -
DR PIR; A48767;
DR InterPro; IPR002129; Pyridoxal dec.
DR Pfam; PF0282; pyridoxal dec; I.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE NEG.
KW Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding.
FT BINDING 277 277 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT DOMAIN 469 500 CALMODULIN-BINDING.
SQ SEQUENCE 500 AA; 56726 MW; 72A043CB885AE10D CRC64;
-----
Query Match 86.0%; Score 2250; DB 1; Length 500;
Best Local Similarity 84.9%; Pred. No. 6.9e-150;
Matches 427; Conservative 39; Mismatches 33; Indels 4; Gaps 2;
-----
Qy 1 MVLSHAVSESDVSHSTFASRYVRTSLPRFKMPENSIPKEAAAYQIINDELMDDGNPRLN 60
Db 1 MVLKTSVSDVSHSTFASRYVRTSLPRFKMPDPSIPKEAAAYQIINDELMDDGNPRLN 60
Qy 61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPTTELQRCVNMIAHLFNAPLEEAETA 120
Db 61 ASFVTTWMEPECDKLIMDSINKNYVDMDEYPTTELQRCVNMIAHLFNAPLEGEETA 120
Qy 121 VGTGSSSEAIMLAGLAFKRWQKRAEGKPVDPKNIVTGANVQVCWEKFPARYFEVELKE 180
Db 121 VGTGSSSEAIMLAGLAFKRWQKRAEGKPVDPKNIVTGANVQVCWEKFPARYFEVELKE 180
Qy 181 VKLSEGVYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKKTG 240
Db 181 VKLSEGVYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVEKNKKTG 240
Qy 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVWRNKEDLPE 300
Db 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVWRNKD 300
Qy 301 IPHINYLADQPTFTLNFSGSSQVIAQYQIIRIGHEGYRNVNENKRENMIIVREGLE 360
Db 301 IPHINYLADQPTFTLNFSGSSQVIAQYQIIRIGHEGYRNVNENKRENMIIVREGLE 360
Qy 361 TERFNIVSKDEGVPVAFSLKDSCHTEFEISDMLRRYGMWIPAYTMPNQAQHTVLR 420
Db 361 TERFNIVSKDEGVPVAFSLKDSCHTEFEISDMLRRYGMWIPAYTMPNQAQHTVLR 420
Qy 421 IREDFSRLEAERLVIDEIKWRELDLPSRVIHKISLGQEKSSNSDNLMTVYKSDIDK 480
Db 421 IREDFSRLEAERLVRDIEKVLHELDLTPARVNAKVAEQAANGSEVH---KKTSE 477
Qy 481 QRDIIIGWKKFVAD-RKKTSGIC 502
Db 478 QLEMITAWKKFVEEKKTNRVC 500
-----

```

```

R EMBL; U10034; AAA93132.1; -
R EMBL; AB005238; BAB10520.1; -
R EMBL; AY094464; AAM19834.1; -
R InterPro; IPR002129; Pyridoxal dec.
R Pfam; PF0282; pyridoxal dec; I.
R PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE NEG.
W Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding;
M Multigene family.
I BINDING 277 277 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
I DOMAIN 469 500 CALMODULIN-BINDING.
I CONFLICT 208 208 A -> D (IN REF. 1).
I Q SEQUENCE 502 AA; 57066 MW; 4B8141FF523E0B22 CRC64;
-----
Query Match 99.7%; Score 2607; DB 1; Length 502;
Best Local Similarity 99.8%; Pred. No. 7.8e-175;
Matches 501; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
-----
y 1 MVLSHAVSESDVSHSTFASRYVRTSLPRFKMPENSIPKEAAAYQIINDELMDDGNPRLN 60
b 1 MVLKTSVSDVSHSTFASRYVRTSLPRFKMPDPSIPKEAAAYQIINDELMDDGNPRLN 60
y 61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPTTELQRCVNMIAHLFNAPLEEAETA 120
b 61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPTTELQRCVNMIAHLFNAPLEEAETA 120
y 121 VGTGSSSEAIMLAGLAFKRWQKRAEGKPVDPKNIVTGANVQVCWEKFPARYFEVELKE 180
b 121 VGTGSSSEAIMLAGLAFKRWQKRAEGKPVDPKNIVTGANVQVCWEKFPARYFEVELKE 180
y 181 VKLSEGVYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKKTG 240
b 181 VKLSEGVYVMDPQQAVDMVDENTICVAAILGSTLNGEFEDVKLLNDLLVEKNKKTG 240
y 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVWRNKEDLPE 300
b 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVWRNKEDLPE 300
y 301 IPHINYLADQPTFTLNFSGSSQVIAQYQIIRIGHEGYRNVNENKRENMIIVREGLE 360
b 301 IPHINYLADQPTFTLNFSGSSQVIAQYQIIRIGHEGYRNVNENKRENMIIVREGLE 360
y 361 TERFNIVSKDEGVPVAFSLKDSCHTEFEISDMLRRYGMWIPAYTMPNQAQHTVLR 420
b 361 TERFNIVSKDEGVPVAFSLKDSCHTEFEISDMLRRYGMWIPAYTMPNQAQHTVLR 420
y 421 IREDFSRLEAERLVIDEIKWRELDLPSRVIHKISLGQEKSSNSDNLMTVYKSDIDK 480
b 421 IREDFSRLEAERLVIDEIKWRELDLPSRVIHKISLGQEKSSNSDNLMTVYKSDIDK 480
y 481 QRDIIIGWKKFVAD-RKKTSGIC 502
b 481 QRDIIIGWKKFVAD-RKKTSGIC 502
-----

```

RESULT 2

```

CE_PETHY STANDARD; PRY; 500 AA.
D DCE_PETHY STANDARD; PRY; 500 AA.
C Q07346;
T 01-NOV-1995 (Rel. 32, Created)
T 01-NOV-1995 (Rel. 32, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
E Glutamate decarboxylase (EC 4.1.1.15) (GAD).
AN GAD.
AN Petunia hybrida (Petunia).
AN Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AN Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
AN Asteridae; Lamiales; Solanales; Solanaceae;
AN NCBI_TaxID=4102;
AN [1]
AN SEQUENCE FROM N.A.
AN TISSUE=Petal;
AN MEDLINE=93374956; PubMed=9366104;
AN Baum G.; Chen Y.; Arazi T.; Takatsuji H.; Fromm H.;

```

```

DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding;
KW Multigene family. 276 PYRIDOXAL_PHOSPHATE (BY SIMILARITY).
FT BINDING 276 494 CALMODULIN-BINDING (BY SIMILARITY).
FT DOMAIN 461 494 CALMODULIN-BINDING (BY SIMILARITY).
SQ SEQUENCE 494 AA; 56140 MW; 741E83A25DC8C48C CRC64;
Query Match 81.5%; Score 2130; DB 1; Length 494;
Best Local Similarity 80.2%; Pred. No. 1.7e-141;
Matches 404; Conservative 46; Mismatches 42; Indels 12; Gaps 3;
QY 1 MVLSHAVSESDVSVHSTPASRYVTFSLPRFKVQPNVSSIPKGAAYQIINDELMDGNPRLNL 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MVLTKTAT-NDSEVCTMFGSRYVTRTLPKVEIGENSIPKQAAAYQIINKDELMDGNPRLNL 59
QY 61 ASFVVTWMEPECDKLMSSINKNVMVDYVPTTELQNRNCVNMIAHLFNAPLEEAETAVG 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
60 ASFVVTWMEPECDKLMSSINKNVMVDYVPTTELQNRNCVNIIFARNAPLEEAETAVG 119
QY 121 VGTVGSSEAIMLAGLAFKRQWKQKKAEGKPFVDPKPNVITGANVQCWKEKPFARYFEVLSKE 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
120 VGTVGSSEAIMLAGLAFKRQWKQKKAEGKPFVDPKPNVITGANVQCWKEKPFARYFEVLSKE 179
QY 181 VKLSEGYVMDPQAAVMVDENTICVADILGSTLNGEFGEDVKLNDLLLVKKNKTGDWTF 240
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 VNLSEGYVMDPDKAAEMVDEITICVAAIIGSTLNGEFGEDVKRLNDLLLVKKNKNEETGWNTP 239
QY 241 IHVDAASGFIAPFLYPELWDFRLPLVKSINVSHKIVLVYAGIVGVIWENKEDLPEEL 300
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
240 IHVDAASGFIAPFLYPELWDFRLPLVKSINVSHKIVLVYAGIVGVIWENKEDLPEEL 299
QY 301 IFHNYLGADQPTFLNFSKGSQVIAQYQYQIIRLHGHEGRVNMVNCRENNIIVLREGLEK 360
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 IFHNYLGADQPTFLNFSKGSQVIAQYQYQIIRLHGHEGRVNMVNCRENNIIVLREGLEK 359
QY 361 TERFNIVSKDQVPLVAFSLKDSCHTFEFISDLRRYGVITPAYTMPNAAHQHITVLRV 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
360 TERFNIVSKDQVPLVAFSLKDSCHTFEFISDLRRYGVITPAYTMPNAAHQHITVLRV 419
QY 421 REDFSRLAERLVDIEKVMREDELPSRVVIHKISLGQEKSESNLMLVTVKKSDDTK 480
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
420 REDFSRLAERLVDIEKVMREDELPSRVVIHKISLGQEKSESNLMLVTVKKSDDTK 479
QY 481 Q--RDIITGWKKFVADRKKTSGIC 502
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
471 EILMEVIVGWRKFKVKRKKMGVC 494
RESULT 4
DCE LYCES STANDARD; PRT; 502 AA.
ID DCE LYCES STANDARD; PRT; 502 AA.
AC P54767;
DT 01-OCT-1996 (Rel. 34; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Glutamate decarboxylase (EC 4.1.1.15) (GAD) (ERT D1) .
OS Lycopersicon esculentum (Tomato) .
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Solanum.
OC NCBI_TaxID=4081;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Ailisa Craig; TISSUE=Pericarp;
RX MEDLINE=95284363; PubMed=7766895;
RA Gallego P.P., Whotton L., Pictor S., Grierson D., Gray J.E.;
RT "A role for Glutamate decarboxylase during tomato ripening: the
RT characterisation of a cDNA encoding a putative glutamate
RL decarboxylase with a calmodulin-binding site.";
RL Plant Mol. Biol. 27:1143-1151(1995).
CC !- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING

```

```

CE2 ARATH STANDARD; PRT; 494 AA.
D DCE2 ARATH STANDARD; PRT; 494 AA.
C Q42472;
T 01-NOV-1997 (Rel. 35; Created)
T 01-NOV-1997 (Rel. 35; Last sequence update)
T 28-FEB-2003 (Rel. 41; Last annotation update)
E Glutamate decarboxylase 2 (EC 4.1.1.15) (GAD 2) .
S GAD2 OR GDH2 OR ATIG65960 OR F12P19.12.
N Arabidopsis thaliana (Mouse-ear cress).
C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
C eurosids_II; Brassicales; Brassicaceae; Arabidopsis.
X NCBI_TaxID=3702;
X [1]
P SEQUENCE FROM N.A.
P STRAIN=cv. Columbia;
P MEDLINE=98363649; PubMed=9700069;
X A Zak M., Arai T., Snedden W.A., Fromm H.;
T "Two isoforms of glutamate decarboxylase in Arabidopsis are regulated
T by calcium/calmodulin and differ in organ distribution.";
L Plant Mol. Biol. 37:967-975(1998).
L [2]
P SEQUENCE FROM N.A.
P STRAIN=cv. Columbia;
P MEDLINE=97267149; PubMed=9112779;
X A Turano F.J., Thakkar S.S., Fang T., Weisemann J.M.;
T "Characterization and expression of NAD(P)-dependent glutamate
T dehydrogenase genes in Arabidopsis.";
L Plant Physiol. 113:1329-1341(1997).
L [3]
P SEQUENCE FROM N.A.
P STRAIN=cv. Columbia;
P MEDLINE=21016719; PubMed=11130712;
X A Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
A White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
A Buehler E., Chan A., Chao K., Chen H., Cheuk R.F., Chin C.W.,
A Chung M.K., Conn L., Conway A.R., Creasy T.H., Dewar K.,
A Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
A Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
A Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
A Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
A Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
A Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
A Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
A Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
A Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
A Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
A Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
A Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
T "Sequence and analysis of chromosome 1 of the plant Arabidopsis
T thaliana.";
L Nature 408:816-820(2000).
C !- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING
C IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS WAY, DIRECTLY
C OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA
C BIOSYNTHESIS (BY SIMILARITY).
C !- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2) .
C !- COFACTOR: Pyridoxal phosphate.
C !- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
C GAD, HDC AND TVRDC).
C ---
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@sib-sib.ch).
C -----
C R EMBL; U49937; AAC31617.1; --
C R EMBL; U46665; AAC33485.1; --
C R EMBL; AC009513; AAF06056.1; --
C R PIR; H96683; H96683.

```

IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS MAY, DIRECTLY OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA BIOSYNTHESIS (BY SIMILARITY).
 -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanate + CO(2).
 -!- COFACTOR: Pyridoxal phosphate.
 -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC, GAD, HDC AND TYRDC).

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; X80840; CAAS56812.1; -;
 PIR; S56177; S56177.
 InterPro; IPR002129; Pyridoxal_dec.
 Pfam; PF00282; pyridoxal_dec; 1.
 PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
 Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding.
 BINDING 278 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 DOMAIN 471 502 CALMODULIN-BINDING (BY SIMILARITY).
 SEQUENCE 502 AA; 56785 MW; 1C5F9ED0084272A6 CRC64;

Query Match 77.0%; Score 2014; DB 1; Length 502;
 Best Local Similarity 76.0%; Pred. No. 2.2e-133;
 Matches 387; Conservative 53; Mismatches 55; Indels 14; Gaps 5;

1 MVL-SHAVESDVSVHSTFASRVVRLSPREFKMPENSIPEKAAYQIINDELMLDGNPRLN 59
 1 MVLTTISIRDSSELSHCTFASRVVRLSPREFKMPENSIPEKAAYQIINDELMLDGNPRLN 60
 60 LASFVTTWMEPECKLIMSINKNKYVDMDEYPTTTELQNRQVNMIAHLFNAPLEAEATAV 119
 61 LASFVSTWMEPECKLIMSINKNKYVDMDEYPTTTELQNRQVNMIAHLFNAPLEAEATAV 120
 120 GVGTVGSSBAIMLAGLAFKRWQSKKAEKPKVDPKNIIVTGANVQVCWEKFKARYFEVELK 179
 121 GVGTVGSSBAIMLAGLAFKRWQSKKAEKPKVDPKNIIVTGANVQVCWEKFKARYFEVELK 180
 180 EVKLSEGYVMPDQAVDMDVENTICVADILGSLNGEFEDVKLLNDLVEKNKGTGWDT 239
 181 EVKLSEGYVMPDQAVDMDVENTICVADILGSLNGEFEDVKLLNDLVEKNKGTGWDT 240
 240 PIHVDAASGGFIAPFLYPLELWDFRPLVKSINVSCHKYGLVYAGIYWRNKEDLPEE 299
 241 PIHVDAASGGFIAPFLYPLELWDFRPLVKSINVSCHKYGLVYAGIYWRNKEDLPEE 300
 300 LIIPHINYLGAQDPTFLNFSKSSQVIAQYQLIRLGHGEGYRVNENKRNMIIVLRGLE 359
 301 LVPHINYLGSQDPTFLNFSKSSQVIAQYQLIRLGHGEGYRVNENKRNMIIVLRGLE 360
 360 KTERFNIVKDEGVPLVAFSLKSSCHTFEFTSDMLRRYGVIVPVTMPPNAQHITVLRV 419
 361 KMGREDIVSKDVGVPVAFSLRDSKSKYTFVSEHLRFGVIVPVTMPPNAQHITVLRV 420
 420 VIREDFSRILAERLVLDIEKVMRELPFSLVTHKI--SLGQKSESNSDNI---MVTV 473
 421 VIREDFSRILAERLVLDIEKVMRELPFSLVTHKI--SLGQKSESNSDNI---MVTV 480
 474 KKSDDIKQDRIITGWKFKVADRKKTSGIC 502
 481 -----ETQKDIKHKRWKRIAG--KKTSGVC 502

RESULT 5
 CE_LACIA STANDARD; PRT; 466 AA.
 D DCE_LACIA STANDARD; PRT; 466 AA.
 C O9CG20; O50645;
 T 16-OCT-2001 (Rel. 40, Created)
 T 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glutamate decarboxylase (EC 4.1.1.15) (GAD).
 GN GADB OR LL1290.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_taxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 1-15, AND CHARACTERIZATION.
 RC STRAIN=01-7;
 RX MEDLINE=99337071; PubMed=10411264;
 RA Nomura M., Nakajima I., Fujita Y., Kobayashi M., Kimoto H., Suzuki I., Aso H.;
 RT "Lactococcus lactis contains only one glutamate decarboxylase gene."
 RL Microbiology 145:1375-1380(1999).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K., Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 CC -!- FUNCTION: CONVERTS INTERNALIZED GLUTAMATE TO GABA AND INCREASES THE INTERNAL PH. INVOLVED IN GLUTAMATE-DEPENDENT ACID RESISTANCE (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanate + CO(2).
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- MISCELLANEOUS: The enzyme is maximally active at pH 4.7. The activity is stable at acidic pH values; there is no activity in the neutral pH range. At pH 4.1 the enzyme activity iss retained at temperatures up to 70 degrees Celsius.
 CC -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC, GAD, HDC AND TYRDC).

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AB010789; BAA24585.1; -;
 PIR; B86786; B86786.
 InterPro; IPR002129; Pyridoxal_dec.
 Pfam; PF00282; pyridoxal_dec; 1.
 PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
 Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
 BINDING 277 277 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SEQUENCE 466 AA; 53926 MW; BCDC732253E473C2 CRC64;

Query Match 39.3%; Score 1028; DB 1; Length 466;
 Best Local Similarity 45.0%; Pred. No. 1.6e-64;
 Matches 195; Conservative 91; Mismatches 141; Indels 6; Gaps 3;

16 FASRYVTSLPRFKMPENSIPEKAAYQIINDELMLDGNPRLNLAASFTVTTWMEPECKLIM 77
 18 FGSSEQVLDLKYKLAQOSIEPRVAYQLVQDMDLDEGNARLNLAATFCQTYMEPEAVKLS 77
 78 SSINKNYVDMDEYPTTTELQNRQVNMIAHLFNAPLEAEATAVGVTVGSSEAIMLAGLAF 137
 78 QTLKNAIDKSEYPRTTTEIENRCVNMIAHLWNA--SEKFKFMGTSTIGSSEACMLGGWAM 135
 138 KRQWNRKAEKPKVDP--KPNIVTGANVQVCWEKFKARYFEVELKVEKLSSEGYVMDQQA 195
 136 KFSWRKRAKLGLDINAKKPNLVISSSGYQVCWEKFCYIWDIEMREVPMDEKHMSSINDKV 195
 196 VDMVDENTICVADILGSLNGEFEDVKLLNDLVEKNKGTGWDTPIHVDAASGGFIAPFL 255
 196 MDYVDEYTIIGVVGIMGITYTGRYDDIKALDNLIEYNKQTDYKVIHVDAASGGLYAPFV 255

SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
MEDLINE=21074935; PubMed=11206551;
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Fosfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
Nature 409:529-533(2001).
[5]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
MEDLINE=21156231; PubMed=11258796;
Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
Lida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kunara S., Shiba T., Hattori M., Shingawa H.;
"Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12";
DNA Res. 8:11-22(2001).
[6]
SEQUENCE OF 1-318 FROM N.A.
SPECIES=E.coli; STRAIN=K12;
Turilin E., Gasser F., Biville F.;
"Sequence and functional analysis of an Escherichia coli DNA fragment
able to complement pqgE and pqgF from Methylobacterium organophilum";
Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE OF 1-15.
SPECIES=E.coli; STRAIN=K12;
MEDLINE=93204884; PubMed=8455549;
Yoshida T., Veguchi C., Yamada H., Mizuno T.;
"Function of the Escherichia coli nucleoid protein, H-NS: molecular
analysis of a subset of proteins whose expression is enhanced in a
hns deletion mutant.";
Mol. Gen. Genet. 237:113-122 (1993).
[8]
SEQUENCE FROM N.A.
SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
MEDLINE=22272406; PubMed=12384590;
Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Sun J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Xiang L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
"Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157";
Nucleic Acids Res. 30:4432-4441(2002).
C -1- FUNCTION: CATALYZES THE PRODUCTION OF GABA.
C -1- CATALYTIC ACTIVITY: L-Glutamate = 4-aminobutanoate + CO(2).
C -1- COFACTOR: Pyridoxal phosphate.
C -1- SUBUNIT: Homohexamer.
C -1- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
GAD, HDC AND TYRDC).

This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
of the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M84025; AAA23834.1; -;
EMBL; AE000246; AAC74566.1; -;
EMBL; D90791; BAA15163.1; -;
EMBL; D90790; BAA15157.1; -;
EMBL; AE000356; AAG56275.1; -;
EMBL; AP002557; BAB35521.1; -;
EMBL; X71917; CAA50736.1; ALT SEQ.
EMBL; AE015194; AAM43309.1; ALT_INIT.

DR PIR; B43332; B43332.
DR PIR; B90891; B90891.
DR PIR; G85726; G85726.
DR EcoGene; EG11490; gadB.
DR InterPro; IPR002129; Pyridoxal.deC.
DR Pfam; PF00282; Pyridoxal.deC.1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family;
KW Complete proteome.
FT BINDING 276 276 PYRIDOXAL PHOSPHATE.
SQ SEQUENCE 466 AA; 52668 MW; 8E6653330A3CSBAED CRC64;
[5]
Query Match 38.7%; Score 1013; DB 1; Length 466;
Best Local Similarity 45.0%; Pred. No. 1.8e-63;
Matches 197; Conservative 83; Mismatches 152; Indels 6; Gaps 3;
OY 16 STEASRYVRSLSRFLKMPENSIPKAAAYQIINDELMLDGNPRNLASVFTVWVPEPCDKL 75
Db 16 SRFAKSIISTAESKRFPLHEMRDDVAFQIINDELVDGNARQNLATFCOTWDDENVHKL 75
OY 76 IMSSINQVYDMDEYPTTELQNRQVNMIAHLFNAPLEEAETAVGVGTGSSSEAIMLAGL 135
Db 76 MDLSINKNWIDKBEYPOSAALDLRCVNWVADLWHPAPKNGQAVGTWTIGSSSEACMLGGM 135
OY 136 AFKRWQNKRAEKGKPVDKENIVTCANVQVCEKFAFYFVELKEVKLSEGYVMDPQQA 195
Db 136 AMKWRWRKREAPAGKPTDKPNLVCGP-VQICWEKFAFYWDVLELRPELPMRFGQFLMPPKRM 194
OY 196 VDMVDENTICVADILGSLNGEPEDVKLLNDLLEKKNKGTGWDTPHVDAAASGGFIAPFL 255
Db 195 IEACDENTIGVVFYGVYTGVEFFQPLHDALDKFOADTGIDIDHIDHIDAAASGGFLAPFV 254
OY 256 YPELEWDFRPLPVKSIINVSCHKVGLVYAGIWIWRNKEDLPEELPHINYLGADQPTFT 315
Db 255 APDIWDFRPLPVKSIINVSCHKVGLVYAGIWIWRNKEDLPEELPHINYLGADQPTFT 314
OY 316 LNFSGSSQVIAQYQIIRLQHEGYRNVMENMIVREGLEKTERFNIV---SKDEG 372
Db 315 INFSRPAQVIAQYEFRLGREGYTKVQNASYQVAAYLADEIAKLGYPYEFICTGREDEG 374
OY 373 VPLVAFSLKQSS--CHTEFEISDMLRRYGVIVPAVTPPNAQHTVLRVVRDFSRTLA 430
Db 375 IPAVCFKLKQEGDEPCYLYLDSLRLRLRGVQVPAFTLGGSEATDILVWKRIMCRGFEFMDFA 434
OY 431 ERLVIDIEKVMRELDLDP 448
Db 435 ELLLEDYKASLKLSDHP 452
RESULT 9
DCEA_ECO57 STANDARD; PRT; 466 AA.
AC P58228;
DT 16-OCT-2001 (Rel. 40, Created)
DT 18-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 43, Last annotation update)
DE Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).
GN GADA OR GADS OR Z4930 OR ECS4397.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Fosfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RT

182 VPMDKHELSLDVKEVFEVDEYTIIGVILGIIYTKGFKDDIALDLDBKVEAYNEANBHQIV 241

241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSCHKYGLVYAGIOWIWRNKEDLPEEL 300

242 IHIDGASGAMFTFVNPPELWDFRLKNVSVINSGHKYGLVYGVGWILWKOKEYLPEKL 301

301 IFHINYLGADOPTFTLNFSGKSSQVIAQYQLRLGHEGRYRNMENCRNMIVLRGLEK 360

302 IFEVSYLGGSMPTWAINFNSASQIIQQYVNFRLYGFYEGYREHTEKTKTALYLSKTVBK 361

361 TERFNIVSKDEGVLVAFSLKDS--HTEFELSDMLRRYGVIVPAYTWPPNAQHITVLR 418

362 SGFELINDGSLPIVCYKLDLWEDWILYDADQLLWKGQVVPAYLPADLSDTIIR 421

419 VVIREDFRTRLAERLVIDIEKVMRELDLPSRVHRIKISLQEKSES 464

422 FVCRADLGINVABEFAADPADLHNLH--ARVLYH--DKERNDS 462

RESULT 11

CEA LISMO STANDARD; PRT; 462 AA.

C Q9F53; Q8Y9S6;

T 28-FEB-2003 (Rel. 41, Created)

T 28-FEB-2003 (Rel. 41, Last sequence update)

T 28-FEB-2003 (Rel. 41, Last annotation update)

E Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).

N GADA OR LMO0447.

S Listeria monocytogenes.

C Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

X NCBI_TaxID=1639;

X [1]

P SEQUENCE FROM N.A.

N STRAIN=LO28 / Serovar 1/2c;

C MEDLINE=21206233; PubMed=11309128;

X Cotter P.D., Gahan C.G.M., Hill C.;

T "A glutamate decarboxylase system protects *Listeria monocytogenes* in

T gastric fluid.";

T Mol. Microbiol. 40:465-475 (2001).

L [2]

N SEQUENCE FROM N.A.

P STRAIN=EGD-e / Serovar 1/2a;

X MEDLINE=21537279; PubMed=11679669;

X Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

A Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

A Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

A Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

A Etian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

A Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

A Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,

A Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

A Nordtsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,

A Remmel B., Rose M., Schluster T., Simoes N., Tierrez A.,

A Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

T "Comparative genomics of *Listeria* species.";

T Science 294:849-852(2001)

L -!- FUNCTION: Converts internalized glutamate to GABA and increases

C the internal pH. Involved in glutamate-dependent acid resistance

C in gastric fluid.

C -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).

C -!- COFACTOR: Pyridoxal phosphate (By similarity).

C -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,

C GAD, HDC AND TRDC).

C

C This SWISS-PROT entry is copyright. It is produced through a collaboration

C between the Swiss Institute of Bioinformatics and the EMBL outstation -

C the European Bioinformatics Institute. There are no restrictions on its

C use by non-profit institutions as long as its content is in no way

C modified and this statement is not removed. Usage by and for commercial

C entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

C or send an email to license@isb-sib.ch).

C

C EMBL; AF309076; AAG22560.1; -.

DR EMBL; ALS91975; CAC98526.1; -.

DR PIR; AH1130; AH1130.

DR ListiList; LMO00447; -.

DR InterPro; IPR002129; Pyridoxal dec.

DR Pfam; PF00282; pyridoxal dec; I.

DR PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.

KW Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.

FT BINDING 273 273 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

FT CONFLICT 152 152 S -> N (IN REF. 1).

SQ SEQUENCE 462 AA; 52498 MW; EA1A442E3E1CE2FA CRC64;

Query Match 36.5%; Score 955.5; DB 1; Length 462;

Best Local Similarity 41.3%; Pred. No. 1.8e-59;

Matches 183; Conservative 94; Mismatches 157; Indels 9; Gaps 4;

QY 7 VSESDVSVHSTFASRYVRTSLPRFKPENSIPKEAAVQIINDELMLDGNPRLNLASFVTT 66

DB 6 VEQNNVPVFGSFES---GQDLPEKRMKESVDPLAIYQLVKDQLIDGSSARQNLATFCQT 62

QY 67 WMEPFCDKLIMSSINKNYVDMDEYPTTELQNRNVNMIHLEFNAPLEBEAETAAGVGVTVGS 126

DB 63 YMEPEAEQIMAEITMEKNAIDKSEYPTAKLESSCVNMLADLWN--VDESEHYMGSTVGS 120

QY 127 SEATMLAGLAEKRWQNKRAEGKPV--DKENIVTGANVOYCWKEKFARYFVELKEYKLS 184

DB 121 SEACMLGGWAMKFRWRSAAALRNGLDIHAKEPSLVISSGQVCWEKFCVYWDIELREVPMS 180

QY 185 EGYVMVDPQAVDMVDENTICVADILAGSTLNGEPEDYKLLMDLLVEKNKKTGMDTPHVD 244

DB 181 EEHLSINDIIMDYVDEYTIIGVILGIIYTKGFKDDIANTLNDLVEDYNNTHDNEVHIVD 240

QY 245 AASGGFIAPFLYPELEWDFRLPLVKSINVSCHKYGLVYAGIOWIWRNKEDLPEELPHI 304

DB 241 GASGAMFTFVNPPELWDFRLKNVSVINSGHKYGLVYGVGWILWKOKEYLPEELVFDV 300

QY 305 NYLGADQPTFTLNFSGKSSQVIAQYQLRLGHEGRYRNMENCRNMIVLRGLEKTERF 364

DB 301 SYLGGHMTMAINFRSASQIIQQYVNFRLYGFYEGYREYQIHWTRDGLALQLSQVAETGLF 360

QY 365 NIVSKDEGVLVAFSLKDS--CHTEFEISDMLRRYGVIVPAYTWPPNAQHITVLRVIVR 422

DB 361 EYNDGANLPIVCYKLDKDDANVANWTLYDLADLQKMGQVFPAYLPEKEMGNTIQRVCR 420

QY 423 EDFRTRLAERLVIDIEKVMRELD 445

DB 421 GDLGNVVTAFKNDLISEISEBLN 443

RESULT 12

DECB LISMO STANDARD; PRT; 467 AA.

ID DCEC LISMO

AC Q8Y4K4;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probabl glutamate decarboxylase gamma (EC 4.1.1.15) (GAD-gamma).

GN LMO2434.

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI_TaxID=1639;

RN [1]_

RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;

RC MEDLINE=21537279; PubMed=11679669;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordtsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,

Rommel B., Rose M., Schlueter T., Simoes N., Tierrez A., Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.; "Comparative Genomics of *Listeria* species."; Science 294:849-852(2001).

1- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).

1- COFACTOR: Pyridoxal phosphate (By similarity).

1- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC, GAD, HDC AND TYRDC).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL: AL591983; CAD00512.1; --
 PIR: AB1379; AB1379.
 ListList; LMO02434; --
 InterPro; IPR002129; Pyridoxal_deC.
 Pfam; PF00282; pyridoxal_deC; I.
 PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE NEG.
 Lyase; Decarboxylase; Pyridoxal phosphate. Complete proteome.
 BINDING 278 278 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 Q SEQUENCE 467 AA; 53640 MW; 648ACB6FDC82B8CF CRC64;

Query Match 36.5%; Score 953.5; DB 1; Length 467;
 Best Local Similarity 40.4%; Pred. No. 2.6e-59;
 Matches 182; Conservative 96; Mismatches 162; Indels 11; Gaps 5;

Y 18 FASRYVRSLSPPKPKPENSIPKEAAQIINDELMDGNPRLNLAASFVITWMEPCDKLIM 77
 b 19 FGSEESTSPKPKVKEPRPRAYQVLDQMLDGNARONLATFCQYMEKEAEILMA 78
 Y 78 SSINKNVYVMDPEYVITTELRQNCVNMIAHLFNAPLEAEATAVGVTGSSSEAIMLAGLAF 137
 b 79 ETEKNAIDKSEYPTALENRQNVNIADLWNAFKD--MSYLTGTVGSSSEACMLGGLAM 136
 Y 138 KRQWQKRAEKGKVD--KENIVTGANVQVCEKFAFYFEVELKEVKLSGYYVMDPQQA 195
 b 137 KFRWNNAEKRLDIOAKRNLNLISSGYQVCEKFCVYVMDVDRVMPMDKRNHLSLDVQKV 196
 Y 196 VDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKKEKTEGWDTPFHVDAASGGFIAPFL 255
 b 197 FDLVDEYITGVVGLITGYTKGDEIQLLDEKVEAYNETNEHQVLVHIDGASGAMFTPFV 256
 Y 256 YPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEELIFHINYLGAOQPTFT 315
 b 257 NPELPWDFRLKQNVVINTSGHKYGLVYVGGVWILWQKDEYLPKELIFEYVLSGSMPTWA 316
 Y 316 LNFSGSSQVIAQYYQLIRLGHGHRNVMENCRNMIVLREGLEKTERFNVSKDEGVPL 375
 b 317 INFERSASQIGQYVNFRLYGFEGYREIHKTRKTAIYLAKTVKSGYFEIINDGANLPI 376
 Y 376 VAFSLKDSKC--HTEFEISDLRRYGVIPAYWPPNAQHHITVLRVVIREDFSETLAERL 433
 b 377 VCYKKEGLEVWEFLYDLAQLLWKGQVFAYPELPADLSITTIQRFVCRADLGNVAEEF 436
 Y 434 VIDLEKVMREIDELPSRVIRKISLGQKSES 464
 b 437 AADFADAIHNLEH--ARVLYH---DKERNDS 462

RESULT 13
 CEB LISIN STANDARD; PRT; 464 AA.
 D DCEB_LISIN
 C Q928R9;
 T 28-FEB-2003 (Rel. 41, Created)
 T 28-FEB-2003 (Rel. 41, Last sequence update)
 T 28-FEB-2003 (Rel. 41, Last annotation update)
 E Glutamate decarboxylase beta (EC 4.1.1.15) (GAD-beta).
 N GADB OR LIIN2463.

Listeria innocua.
 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Blochrieser C., Rusniok C., Amend A., Baquero F., Berche P., Blocker H., Brandt P., Chakraborty T., Chahbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P., Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O., Entian K.-D., Fshhi H., Garcia-del Portillo F., Garrido P., Gautier L., Geobel U., Gomez-Lopez N., Hain T., Hauf J., Jackson D., Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G., Madheno E., Waitournam A., Mata Vicente J., Ng E., Nedjari H., Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R., Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A., Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.; "Comparative genomics of *Listeria* species."; Science 294:849-852(2001).

1- FUNCTION: Converts internalized glutamate to GABA and increases the internal pH. Involved in glutamate-dependent acid resistance in gastric fluid (By similarity).

1- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).

1- COFACTOR: Pyridoxal phosphate (By similarity).

1- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC, GAD, HDC AND TYRDC).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL: AL596172; CAC97690.1; --
 PIR: AB1740; AB1740.
 ListList; LIN02463; --
 InterPro; IPR002129; Pyridoxal_deC.
 Pfam; PF00282; pyridoxal_deC; I.
 PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE NEG.
 Lyase; Decarboxylase; Pyridoxal phosphate. Complete proteome.
 BINDING 275 275 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 464 AA; 53599 MW; 4C35CD1395ADF481 CRC64;

Query Match 36.3%; Score 948; DB 1; Length 464;
 Best Local Similarity 41.7%; Pred. No. 6.2e-59;
 Matches 180; Conservative 92; Mismatches 154; Indels 6; Gaps 3;

QY 18 FASRYVRSLSPPKPKPENSIPKEAAQIINDELMDGNPRLNLAASFVITWMEPCDKLIM 77
 Db 16 FGSSAEDRDIKPYTLAKEPLPRPRAYRIVKDELDEGSAQONLATFCQYMEDEATKIMS 75
 QY 78 SSINKNVYVMDPEYVITTELRQNCVNMIAHLFNAPLEAEATAVGVTGSSSEAIMLAGLAF 137
 Db 76 ETEKNAIDKSEYPTALENRQNVNIADLWNAFKD--QKFNQSTIGSSSEACMLGGLAM 133
 QY 138 KRQWQKRAEKGKPY--DKPNIVTGANVQVCEKFAFYFEVELKEVKLSGYYVMDPQQA 195
 Db 134 KFAWRKRAEKLGLDIYAQKPNLNISSGYQVCEKFCVYVMDVDRVMPMDKRNHLSLDVQ 193
 QY 196 VDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKKEKTEGWDTPFHVDAASGGFIAPFL 255
 Db 194 LDYVDEYITGVVGLITGYTKGDDIYALNEKLEBEYNSKTDYKVIIVHDAASGFFTPFV 253
 QY 256 YPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEELIFHINYLGAOQPTFT 315
 Db 254 EPDIILWDFRLKQNVVINTSGHKYGLVYVGGVWILWQKDEYLPKELIFEYVLSGSMPTWA 313
 QY 316 LNFSGSSQVIAQYYQLIRLGHGHRNVMENCRNMIVLREGLEKTERFNVSKDEGVPL 375
 Db 314 INFERSASHIIGQYVNFRLYGFEGYREIHKTRKTAIYLAKTVKSGYFEIINDGANLPI 373

376 VAFSLKDS--CHTEFEISDMLRRYGVIVPMTYPPNACHITVLRVWIREDFSRFLAERL 433
 374 VCYKLDANVWKTLYDLADRQMRGQVWPAYPLPKLENLIIQRYVCKADLGFENMABEF 433
 434 VIDIEKVMRELD 445
 434 IQDFQASIQELN 445

RESULT 14
 DEEB LISMO STANDARD; PRT; 464 AA.
 Q9EYW9; Q8Y4S0; Q9AGQ0;
 28-FEB-2003 (Rel. 41, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 Glutamate decarboxylase beta (EC 4.1.1.15) (GAD-beta).
 GADB OR LMO2363.

Listeria monocytogenes.
 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 NCBI_TaxID=1639;
 [1]_SEQUENCE FROM N.A.
 STRAIN=LO28 / Serovar 1/2c, and EGD5;
 MEDLINE=21206233; PubMed=11309128;
 Cotter P.D., Gahan C.G.M., Hill C.;
 "A glutamate decarboxylase system protects *Listeria monocytogenes* in gastric fluid";
 Mol. Microbiol. 40:465-475 (2001).
 [2]

SEQUENCE FROM N.A.
 STRAIN=EGD-e / Serovar 1/2a;
 MEDLINE=21537279; PubMed=11679669;
 Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetoui F., Couve E., de Daruvar A., Delouch P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Krefz J., Kuhn M., Kunst F., Kurapkak G.,
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 "Comparative genomics of *Listeria species.*";
 Science 294:849-852(2001).
 !- FUNCTION: Converts internalized glutamate to GABA and increases the internal pH. Involved in glutamate-dependent acid resistance in gastric fluid.

!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
 !- COFACTOR: Pyridoxal phosphate (By similarity).
 !- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC, GAD, HDC AND TYRDC).

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF309077; AAG22562.1; -
 EMBL; AF329447; AAK17187.1; -
 EMBL; AL591983; CAD00441.1; -
 PIR; AC1370; AC1370.
 ListLiSt; LMO2363; -
 Rfam; PF00282; pyridoxal dec; 1
 PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
 Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
 BINDING 275 275 PYRIDOXAL PHOSPHATE (BY SIMILARITY). T

FT VARIANT 92 124 124 124 A -> P (IN STRAIN LO28).
 FT VARIANT 124 124 124 124 E -> D (IN STRAIN LO28).
 FT VARIANT 261 261 261 261 F -> L (IN STRAIN EGD5).
 FT VARIANT 375 375 375 375 C -> R (IN STRAIN EGD5).
 FT VARIANT 380 381 381 381 DD -> TT (IN STRAIN LO28).
 SQ SEQUENCE 464 AA; 53542 MW; F2E2778CFD1E2C36 CRC64;

Query Match 35.9%; Score 940; DB 1; Length 464;
 Best Local Similarity 41.4%; Pred. No. 2.2e-58;
 Matches 179; Conservative 93; Mismatches 154; Indels 6; Gaps 3;

Qy 18 FASRYVTSLSRFRKMPENSIPEKAAAYQIINDELMLDGNPRNLASFTVWMEPECDKLM 77
 Db 16 FGSAEDRDIKPYTLGKLEPRFAVRLVRLKDELLDEGSAEONLATFCQYMEDEATKLM 75
 Qy 78 SSINKNVDMDEYPTTELOQRVCMIAHLFNAPLEBAETAAGVGVTVGSSEAIMLAGLAF 137
 Db 76 ETLKNAIDKSEYPRTAELNRCVNIADLWHAPKD--QKFMGTSTIGSSEACNLGGMAM 133
 Qy 138 KRKQNRKRAEGKPV--DKPNIVTGANVQVCKEFPARYFEVELKVKLSEGYVMDFQA 195
 Db 134 KFAWRKRAEKGLGLDIYAKKPNLVISSGYQVCKEFCVYWDIMRVVPMDEKHMQLNDOV 193
 Qy 196 VDMYDENTICVADITLGSITLNGEEDVKLNDLLEKNETGWDTPIHVDAASGGFIAPFL 255
 Db 194 LDYVDEYTIQVVGILGITYGRYDDIYALNEKLEENSKTDYKYVHVDAASGGFFPFV 253

Qy 256 YPELEWFRPLVKSINVSCHKYGLVYAGIGWIVRNKEDLPEELIFHINYLGADQPTFT 315
 Db 254 EPDIIMDFRLKNWISINTSGHKYGLVYGGVGLWLMKDESYLPEELIFKVSYLGGEMFTMQ 313
 Qy 316 INFSGSSQVIATQYYQLIRLGHGGRVNMENCRNMTVLREGLEKTERFNIIVSKDEGVPL 375
 Db 314 INFGRSASHIIGQVYFNFLRYGFEGRYTIHQKTSVQAQYLAHAYEQTGYFFDFNDGSHLPI 373
 Qy 376 VAFSLKDS--CHTEFEISDMLRRYGVIVPMTYPPNACHITVLRVWIREDFSRFLAERL 433
 Db 374 VCYKLDANVWKTLYDLADRQMRGQVWPAYPLPKLENLIIQRYVCKADLGFENMABEF 433
 Qy 434 VIDIEKVMRELD 445
 Db 434 IQDFQASIQELN 445

RESULT 15
 ID DCE YEAST STANDARD; PRT; 585 AA.
 AC 004792;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutamate decarboxylase (EC 4.1.1.15) (GAD).
 GN GAD1 OR YMR250W OR YMR920.04.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagals K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
 RA "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome
 RT XIII.";
 RL Nature 387:90-93 (1997).
 CC !- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
 CC !- COFACTOR: Pyridoxal phosphate (By similarity).
 CC !- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC, GAD, HDC AND TYRDC).
 CC -----

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - nucleic search, using frame_plus_p2n model

un on: October 22, 2003, 15:14:55 ; Search time 2444 Seconds

(without alignments)

4992.163 Million cell updates/sec

title: US-10-006-852-2

effect score: 2615

sequence: 1 MWLSHAVSESDVSVHSTFAS.....DIITGKMKFVADRKKTSGIC 502

coring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Zgapop 6.0 , Zgapext 7.0

Delop 6.0 , Delext 7.0

searched: 22781392 seqs, 12152238056 residues

total number of hits satisfying chosen parameters: 45562784

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ommand line parameters:

MODEL=frame+p2n.model -DEV=xlp
Q=/cgm2 1/uspto.spool_p/US10006852/runat 21102003 164104 2461/app query.fasta_1.647
DB=EST -QPWT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cgi -LIST=45
DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
OUTPMT=pto -NORW=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US10006852 @CGN 1 1 3549 @runat 21102003 164104 2461 -NCPU=6 -ICPU=3
NO MMAP -LARGEOBVERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_esttm:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_htc:**
- 9: gb_estl:**
- 10: gb_est2:**
- 11: gb_htc:**
- 12: gb_est3:**
- 13: gb_est4:**
- 14: gb_est5:**
- 15: em_estfum:**
- 16: em_estom:**
- 17: em_gss_hum:**
- 18: em_gss_inw:**
- 19: em_gss_pln:**
- 20: em_gss_vrt:**
- 21: em_gss_fun:**
- 22: em_gss_mam:**
- 23: em_gss_mus:**
- 24: em_gss_pro:**
- 25: em_gss_rod:**
- 26: em_gss_phg:**
- 27: em_gss_vrl:**
- 28: gb_gssI:**

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1622	62.0	1168	11	AY103733	Zea mays
2	1332	50.9	811	14	CB655721	CB655721 OSJNEC09G
3	1281	49.0	821	10	BE040860	BE040860 OFI2G05
4	1279	48.9	837	10	BF275438	BF275438 GA_EB002
5	1275	48.8	819	13	EQ165651	EQ165651 EST611520
6	1225	46.8	775	13	EU026594	EU026594 OHG1JF24
7	1220	46.7	786	14	CA813481	CA813481 CA48LN101
8	1217.5	46.6	823	10	BF263784	BF263784 HV_CEA000
9	1203	46.0	762	13	BQ971999	BQ971999 OHB9G09.Y
10	1196	45.7	803	10	BE054156	BE054156 GA_Ba003
11	1189	45.5	750	13	BQ023643	BQ023643 OHF12C03
12	1187.5	45.4	786	13	BQ027739	BQ027739 OHG7H16.Y
13	1185	45.3	746	13	BQ025264	BQ025264 QHF8J20.Y
14	1185	45.3	746	13	BQ027680	BQ027680 QHG7B13.Y
15	1184	45.3	746	13	BQ025649	BQ025649 OHG10J06
16	1182	45.2	752	13	BQ026039	BQ026039 OHG12C09
17	1182	45.2	769	13	BQ025918	BQ025918 OHG12C07
18	1180	45.1	744	13	BQ025839	BQ025839 OHG11L03
19	1179	45.1	744	13	BQ027771	BQ027771 QHG7K14.Y
20	1178	45.0	744	13	BQ025655	BQ025655 QHG10K09
21	1178	45.0	744	13	BQ025893	BQ025893 QHG1P15
22	1177.5	45.0	808	14	CB893221	CB893221 EST646013
23	1175	44.9	744	13	BQ026030	BQ026030 QHG12N09
24	1175	44.9	745	13	BQ025328	BQ025328 QHF8O15.Y
25	1171	44.8	752	13	BQ841623	BQ841623 WHE2983_B
26	1168	44.7	698	14	CD484071	CD484071 atr01-6ms
27	1168	44.7	746	13	BQ027486	BQ027486 QHG5P23.Y
28	1167	44.6	790	10	BG647367	BG647367 EST508986
29	1164	44.5	744	13	BQ025929	BQ025929 OHG12D11
30	1162	44.4	732	12	BQ026249	BQ026249 OHG14D05
31	1161.5	44.4	781	12	B1308490	B1308490 EST529900
32	1160	44.4	781	13	BQ025900	BQ025900 QHG12A05
33	1159	44.3	733	10	BG594946	BG594946 EST493624
34	1154	44.1	721	13	BQ027749	BQ027749 QHG7I110.Y
35	1154	44.1	746	13	BQ026554	BQ026554 QHG17D12
36	1152	44.1	726	13	BQ024451	BQ024451 QHF3G03.Y
37	1152	44.1	726	13	BQ027466	BQ027466 OHG5N09.Y
38	1152	44.1	742	13	BQ916451	BQ916451 QHB17010
39	1151	44.0	793	14	CB341828	CB341828 CA32EN000
40	1150	44.0	742	13	BQ025333	BQ025333 QHF8P03.Y
41	1146	43.8	727	13	BQ023920	BQ023920 QHF14N08
42	1144	43.7	716	13	BQ027449	BQ027449 OHG5L24.Y
43	1143	43.7	715	13	BQ025894	BQ025894 OHG11P17
44	1142	43.7	721	10	BG589644	BG589644 EST497486
45	1142	43.7	743	13	BQ025776	BQ025776 QHG11F24

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	ALIGNMENTS
AY103733	AY103733	Zea mays	AY103733	AY103733.1	HTC	Zea mays	Zea mays	1168 bp mRNA
		Zea mays		GI:21206811				linear
		Zea mays						HTC 16-OCT-2002
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD						
		clade; Panicoideae; Andropogoneae; Zea.						
		REFERENCE						
		1 (bases 1 to 1168)						

AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
AUTHORS 2 (bases 1 to 1168)
TITLE Coe,E.H.
JOURNAL Direct Submission
AUTHORS Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
TITLE If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; IIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES Location/Qualifiers
 1..1168 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:638900"
 /gb_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
 275 a 332 c 345 g 216 t

BASE COUNT 275 a 332 c 345 g 216 t
ORIGIN
 1 MetValLeuSerHisAlaValSer-----GluSerAspValSerValHisSerThrPhe 18
 107 ATGTTGCTTCGGACGGAGCTCGGGCGGGACGACGACGCGGTGGCTCCACCTTC 166
 19 AlaSerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIlePro 38
 167 GCCTGGCGCTACGTGGCGAGCCCTGCCCGGTACCGGATCCGGGCGGTCCGATCCCG 226
 39 LysGluAlaIaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeu 58
 227 CGGAGGCGCGCAGCAGATCATCAGGACGAGCTCATGTGGAGCGCAACCCGGCCCTG 286
 59 AsnLeuAlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSer 78
 287 AACCTGGCGTCTTCGTCCACCACGTGGATGGAGCCCGAGTGCAGCAAGCTCATATGGAC 346
 79 SerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsn 98
 347 TCCATCAACAGAACTACGTCCAGATCGACATGACGAGTACCCCGTCCACTGAGCTCCAGAAC 406
 99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAla 118
 407 GCCTGGCTAAACATGATGAGCTCACTGTTCAACCGCCGATCAAGAGGAGCAGACAGCT 466
 119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
 467 ATTGGGGTCCGACAGTGGATCTCCGAGAGCGATCATGCTTGCCAGGCTGGGCTTCAAG 526
 139 ArgLysTrpGlnAsnLysArgLysAlaGluLysProValaspLysProAsnIleVal 158

Db 527 AGGAAATGGCAGAACAAAGAGGAGGAGCGGGGAAACCGTGTGACAGCCCAACATCGTC 586
Qy 159 ThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeu 178
Db 587 ACCGGTGCCAAAGTTCAGGTTCTGCTGGGAGAGTTTCAAGGTATTTTGGTGGAGCTG 646
Qy 179 LysGluValLysLeuSerGluGlyTyrTrpValMetAspProGlnGlnAlaValAspMet 198
Db 647 AAGGAGGTTAAGCTATCAGAAGGGTACTACGTCATGGACCCCGTCAAGGGTGTGTGAGATG 706
Qy 199 ValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPhe 218
Db 707 GTGGACGAGAACACTATCTGTGTGCGGGCGATCTGTGGGTGCACTCCACCCGGAGAGTTC 766
Qy 219 GluAspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAsp 238
Db 767 GAGGATGTCAGACAGTTCAGACGACTTCTTACAGAGAAGAAACAAGAACTGGGTGGAC 826
Qy 239 ThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGlu 258
Db 827 GTGCCTATPCCACGTCGACGCGCGGAGGGTTCATPAGCGCCCTTCTCTACCCCTGAG 886
Qy 259 LeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyr 278
Db 887 CTCGAGTGGGACTTCAGCTGCCCTCTGGTGAAGAGCATCAATGTCCGCGGGCACAAGTAC 946
Qy 279 GlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGlu 298
Db 947 GGCCTCGTCTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1005
Qy 299 GluLeuIlePheHisIleAsnTyrIleuGlyAlaAspGlnProThrPheThrLeuAsnPhe 318
Db 1006 GAGCTCATCTTCCATATCAACTACCTGGGGAGCGGACCGACCTACCTCCAGCTCAACTTC 1065
Qy 319 SerLysGlySerSerGlnValIleAlaGlnTyrTrpGlnLeuIleArgLeuGlyHisGlu 338
Db 1066 TCCAAAGGTTCTGTGATGATTCATGGCAGGACTATCAACTCACTTCGCGTAGGCTTCGAG 1125
Qy 339 GlyTyrArgAsnValMetGluAsnCysArgGluAsn 350
Db 1126 GGGTACAGAAACATCATGGAGAACTGCCAGGAGAAC 1161

RESULT 2
 CB655721
LOCUS OSJNEC09G09.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEC09G09 5', mRNA sequence.
ACCESSION CB655721 GI:29659446
VERSION EST.
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 811)
AUTHORS Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
JOURNAL Unpublished
COMMENT Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aaa cga cgg cca ctg
 BACKWARD: gga aac agc tat gac cat g

Plate: 09 row: G column: 09
 Seq primer: gta aaa cga cgg cca gtc.
 Location/Qualifiers
 1. .811
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEC0909"
 /tissue type="Leaf"
 /dev stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEC"
 /note="vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
 ASE COUNT 182 a 236 c 244 g 149 t
 RIGIN

Alignment Scores:
 red. No.: 2.52e-144 Length: 811
 core: 1332.00 Matches: 249
 Percent Similarity: 95.91% Conservative: 9
 Best Local Similarity: 92.57% Mismatches: 11
 Query Match: 50.94% Indels: 0
 Gaps: 14

US-10-006-852-2 (1-502) x CB655721 (1-811)

61 AlaserPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerIle 80
 3 GCCTCGTTCTGCCACCGTGGAGCCCGGAGTGGCAGACTCCAGGCTCCGTC 62
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
 63 AACAAAGAACTACGTCGACATGGACAGTACCCCGTACCACCGAACTCCAGAACCGATGT 122
 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 123 GTGAACATGATGTCACACCTCTCAATGCCTCTAGGGACTCTGAACCGCCGTCGGA 182
 121 ValGlyThrValGlySerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 183 GTCGGCACTGTCGGCTCGTCGAGGCCATCATGCTCCCGGTTGGCTTCAAGAGAGG 242
 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 243 TGGCAGAACCAAGATGAAGGCGAGCCGCAAGCCATGCCAACGCTAACATTTGTCCCGGC 302
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 303 GCCAATGTCCAAAGTTTCTGGGAGAAATTCGGCGATCTTCGAGGTTGAGCTCAAGGAA 362
 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnAlaValAspMetValAsp 200
 363 GTGAGCTGAGTGCAGCGCTACTAGTGCATGGACCCAGCTAAGCCGTTGGATATGTTCCGAC 422
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 423 GAGAACCACTCTGCGTCCGCGCGATCCTCGGTCGACCGCTGAACGGGGAGTTTCGAGGAC 482
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 483 GTGAAGCTGCTCAACGATCTCTCACCAGAAAGAACCGCTGAACAGGCTGGGACAGCCCG 542
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 543 ATCCAGCTGGACGGCGGAGGGGGTTCATCCGCGCTTCTGTATCCCGGAGCTGGAG 602
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 603 TGGGACTCCGGCTGCCCTGTTGAGAGGATCAACGTCGAGCGGGCAAGATGACGGGCTC 662
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuGlu 300

Db 663 GTCTACGGCGGATCGGTTGGTGCATCTGGAGGACCAAGAGGATCTCCCTGAGGAGCTC 722
 QY 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 Db 723 ATCTTCCACATCACTACTCTCGCGCCGACCCAGCCACTTCCACTTCACTTCTCCRAAG 782
 QY 321 GlySerSerGlnValIleAlaGlnTyr 329
 Db 783 GGGTTTCAGCCAGGTCATTCACAGTAT 809

RESULT 3
 BE040860
 LOCUS
 DEFINITION
 BE040860.1 GI:8336367
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 1 (bases 1 to 821)
 Bohmert,H.J.; Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrea
 H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C.,
 Scara,G., Wheeler,M. and Zepeda,G.R.
 Functional Genomics of Plant Stress Tolerance
 Unpublished
 Contact: Michalowski,C.B.
 University of Arizona
 Bio Sciences West room 513, Tucson, AZ 85721, USA
 Tel: 520-621-7982
 Fax: 520-621-1697
 Email: cbm@u.arizona.edu
 An open reading frame exists.
 Insert Length: 1 Std Error: 0.00.
 Location/Qualifiers
 1. .821
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /strain="Porkali"
 /db_xref="taxon:4530"
 /tissue type="entire plant"
 /dev stage="2 weeks"
 /clone lib="Op"
 /note="1 week 150mM NaCl"
 BASE COUNT 185 a 231 c 254 g 149 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2.26e-138 Length: 821
 Score: 1281.00 Matches: 243
 Percent Similarity: 94.42% Conservative: 11
 Best Local Similarity: 90.33% Mismatches: 15
 Query Match: 48.99% Indels: 1
 Gaps: 0

US-10-006-852-2 (1-502) x BE040860 (1-821)

QY 46 IleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAlaSerPheValThr 65
 Db 9 ATCAAGCAGGCTGATGTGGACGCAACCCGGGCTGAACCTCGCTCGTTCCTCACC 68
 QY 66 ThrTrpMetGluProGluCysAspLysLeuIleMetSerIleAsnLysAsnTyrVal 85
 Db 69 ACGTGTATGGAGCCCGAGTGGACAGCTCATCCAGGCTTCCAGGAGCTGTGACATGATGCA 128
 QY 86 AspMetAspGluTyrProValThrGluLeuGlnAsnArgCysValAsnMetIleAla 105
 Db 129 GACATGGACGAGTACCCCGTCCAGCCAGCTCCAGAACCGATGTGACATGATGCA 188

106 HisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrValGly 125
 189 CACCTCTTAATGCTCTAGGGGACTCTGAACCGCGCTGGAGTCCGACATGTCGGC 248
 126 SerSerGluAlaLeuMetLeuAlaGlyLeuAlaPheLeuArgLysTrpGlnAsnLysArg 145
 249 TCCTCTGAGGCCATCAGCTCGCCGGTTGGCTTCAAGAGGAGTGGGCAAGCAAGATG 308
 146 LysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGlnVal 165
 309 AAGCGACCGGCAAGCCATCGACAGCCCTAACATGTCACCGCGCCCAATGTCGAAGTT 368
 166 CysTrpGlnLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSerGlu 185
 369 TGCTGGGAGAAGTTCGGCGGATACTTCAGGTTGAGCTCAAGGAAGTGAAGCTGAGTGAC 428
 186 GlyTyrThrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIleCys 205
 429 GGTACTACGTATGAGCCAGCTAAGCCCGTGGATATGTCACCGAGAACACCACTGTC 488
 206 ValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspValLysLeuLeuAsn 225
 489 GTCGCGCGATCTCGGGTCGACGCTGAACGGGGAGTTCGAGGACGCTGAAGCTGCTCAAC 548
 226 AspLeuLeuValGluLysAsnLysGluThrGlyTyrPheAspThrProIleHisValAspAla 245
 549 GATCTGCTCAACCAAGAGAACCGTGAACAGCGTGGGACACCGCGCATCCAGTGGACGG 608
 246 AlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTyrPheArgLeu 265
 609 GCAGCGCGGGTTCATCGCCCGTTCCTGTACCCGAGCTGGAGTGGGACTTTCGGCTG 668
 266 ProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIle 285
 669 CCCTGCTGAAGAGCATCAACGTGAGGGGCGACAAAGTACGGGCTGCTACGCGGGATT 728
 286 GlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuLeuIlePheHisIleAsn 305
 729 CGGTGTGCATCTGGGAGCAAGAGGATCTCCCTGAGGAGCTCATCTTTCACATAACT 788
 306 TyrLeuGluAlaAspGlnProThrPhe 314
 789 TACT--GNGCCGACAGCCACTTTTAC 813

BF275438 837 bp mRNA linear EST 07-MAR-2001
 GA_Eb0024A19f Gossypium arboreum 7-10 dpa fiber library Gossypium
 arboreum cDNA clone GA_Eb0024A19f, mRNA sequence.

BF275438 GI:11206508
 BF275438.1

Gossypium arboreum
 Gossypium arboreum
 Gossypium arboreum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids
 ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 837)
 Wang, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
 D., Wood, T.C., Leslie, A. and Wilkins, T.A.
 An integrated analysis of the genetics, development, and evolution
 of the cotton fiber

Unpublished
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAATACGACTCATATAGGG
 High quality sequence start: 2
 High quality sequence stop: 718.

Location/Qualifiers
 1..837
 /organism="Gossypium arboreum"
 /mol_type="mRNA"
 /strain="AKA"
 /cultivar="8400"
 /db_xref="taxon:29729"
 /clone="GA_Eb0024A19f"
 /tissue_type="Fibers isolated from bolls harvested 7-10
 dpa"
 /lab_host="G. coli"
 /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
 /note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 244 a 171 c 198 g 223 t 1 others

ORIGIN

Alignment Scores:
 Pred. No.: 3,98e-138 Length: 837
 Score: 1279.00 Matches: 248
 Percent Similarity: 94.80% Conservative: 7
 Best Local Similarity: 92.19% Mismatches: 14
 Query Match: 48.91% Indels: 1
 DB: 10 Gaps: 0

US-10-006-852-2 (1-502) x BF275438 (1-837)

QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 32 ATGGTGTATCTAAGACAGCATCTGAAACTGATGCTCAGTTCACTCCACTTTGGCTCT 91
 QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 92 CGTTATGTCGGAACCTCACTGCCAGGTTCAAAATGCCAGAAAACCTCCATACCAAAAGAG 151
 QY 41 AlaAlaTyrGlnIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 Db 152 GCTGTTATCAGATCATCAATGATGACTGATGCTTGTATGGCAACCCAAAGTTGACCTT 211
 QY 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 212 GCCTCTTTTGTACTACATGGATGGAGCTGAATGCTGATAAGCTTATAATGGACTCCATC 271
 QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 Db 272 AACAAAGAACTACGTTGACATGAAATACCCCTGCACCCTTCAGACTTCAGAAATCGTTGT 331
 QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
 Db 332 GTGAACATGATAGCACATCTTTTCAACCGCACCATTTGGGAGAGTCAGAGGCAGCAGTTGGT 391
 QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 392 GTGGGAACAGATTTGGGTCATCGGAGGCAATCATGTTGGCTGGCCCTTTCATCAAGAGAAA 451
 QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 Db 452 TGGCAAAAACACCGTAAAGCTGAAGCAAAACCTTATGAAAAGCCGAAATATCGTGTGGA 511
 QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 512 GCCAATGTTGAGTTGCTGGGAGAAAATTTGCAAGGTACTTTTGAAGTGGAGTTGAAAGAA 571
 QY 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 Db 572 GTGAAGCTTAGAGAAGGGTACTATGATGGACCCTGTCAAAGCAGTGGAAATGGTTGAT 631
 QY 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 Db 632 GAAAAACCACTCTGTGTAGCTCAATCTTGGGTTCAACCCCTCAATCGGAGAAATTTGAAAGAT 691
 QY 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrPro 240
 Db 692 GTCAAGCTTCTAAGCACTTTTATTTGAAAAGACANA-GAAAACCTGGATGGATACCCCC 750

241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 751 ATTCATGTTGATCAGCTAGTGTGGATTATTGGCCCAATTTTGTACCAGAACTTGAG 810
 261 TrpAspPheArgLeuProLeuValLys 269
 b 811 TGGGACTTCAAGGCTCCCTGTGAAG 837

EST 25-APR-2002
 EST 1520 KVVC Medicago truncatula cDNA clone pkVKC-10G4, mRNA
 BQ165651 819 bp mRNA linear EST 25-APR-2002
 sequence.
 BQ165651
 BQ165651.1 GI:20308272
 EST.
 Medicago truncatula (barrel medic)
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 819)
 VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S.,
 Uterback,T., Cheung,F. and Fraser,C.M.
 The Medicago truncatula 'killeclone' set; ESTs selected and
 re-arrayed from various libraries
 Unpublished
 Contact: kvanden@cbs.umn.edu
 Department of Plant Biology
 University of Minnesota
 220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55109, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvanden@cbs.umn.edu
 TIGR sequence name: MTNAJ76TK Alias Clone name:MHAM-5712 More
 information is available at: www.medicago.org
 Seq primer: SKmod (CTA GAA CTA gTg 9AT CC).
 Location/Qualifiers
 1. .819
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="pkVKC-10G4"
 /tissue types="mixed tissues"
 /dev stage="various stages"
 /lab_host="XL0LR"
 /clone_lib="KVVC"
 /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
 was directionally ligated into the Unizap XR vector from
 Stratagene and packaged using Gigapack III Gold packaging
 extracts. plasmids containing cDNA inserts were excised
 from the recombinant lambda-zap phage using Ex-assist
 helper phage and propagated in XL0LR cells."

245 a 158 c 188 g 228 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.12e-137 Length: 819
 Score: 1275.00 Matches: 245
 Percent Similarity: 95.44% Conservative: 6
 Best Local Similarity: 93.16% Mismatches: 11
 Query Match: 48.76% Indels: 1
 DB: 13 Gaps: 0

QY 22 TyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAla 41
 Db 92 TATGTCAGAACTTCACCTCCCTAGATTCAAGATGCCAGAGGAGTCTATACCAAGGATGCA 151
 QY 42 AlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
 Db 152 GCATACCACAAATAATAAAGCATGAATGATGCTTGTATGATGAAAACCCCTAGATTGAATTTGGCA 211
 QY 62 SerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerIleAsn 81
 Db 212 TCCYTTGTGACACACATGATGAGCCCTGAGTGTGATAAACAATCAATCATCTGGCTTCCATTAAC 271
 QY 82 LysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysVal 101
 Db 272 AAGAACTACGTTGACATGGATGAATGCCAGTACCACCTGAGCTACAGAAATCCGCTGTGT 331
 QY 102 AsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGlyVal 121
 Db 332 AACATGAT 391
 QY 122 GlyThrValGlySerSerGlyLysAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrp 141
 Db 392 GGCACCTGTGGCTCATCAGAGGCTATAATGTTAGCTGGATTGGCATTCAAAGGAGTGG 451
 QY 142 GlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAla 161
 Db 452 CAGAAACAAAGAAAACAAAGAGGAAAAGCCCTTATGACAAAACCTAAACATTTGCTCCTGGAGCC 511
 QY 162 AsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluVal 181
 Db 512 AATGTCAGGTTTGTGGGAGAAATTCCAAGGATCTTTGAAATGGAAATTCGAAAGAGGTG 571
 QY 182 LysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGlu 201
 Db 572 AAGCTAAGTGAAGGATACTATGTAATGGACCCCTCAAAAAGCTGTGGAATTCGTAGATGAG 631
 QY 202 AsnThr-IleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspVa 221
 Db 632 AACCCCAATTTGTGTGCTTCCCTTCCCTTCCACACATAAATGGAGAGTTCGAAGATGT 691
 QY 221 LysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrProIle 241
 Db 692 TAAACGCTTAAACAGATCCTTGTGTAAGAAAGAAATGAGAAACTGGATGGGACACTCTTAT 751
 QY 241 eHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTr 261
 Db 752 TCATGTGGATGCAGCTAGTGTGGCTTCATTTGCCATTCATTTATTTATTTATCCAGAACTTGAGTG 811
 QY 261 pAspPhe 263
 Db 812 GGATTTTC 818
 RESULT 6
 BU026594
 LOCUS BU026594
 DEFINITION 775 bp mRNA linear EST 23-AUG-2002
 Clone OHGI7F24, mRNA sequence.
 BU026594
 BU026594.1 GI:22462114
 EST.
 Helianthus annuus (common sunflower)
 Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroidae;
 Heliantheae; Helianthus.
 1 (bases 1 to 775)
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riessberg,L.,
 Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,
 P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
 Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project

REFERENCE
 AUTHORS
 TITLE

http://compenomics.ucdavis.edu/
 Unpublished
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R. W. Michelmore Lab
 University of California at Davis (UCD)
 Asmudson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atc.org [michelmoreveg@mail.ucdavis.edu]
 belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: QHG17 row: F column: 24.

FEATURES

Location/Qualifiers
 1..775
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clone="QH17F24"
 /lab_host="E.coli"
 /clone_lib="QH_EFGHU sunflower RHA280"
 /notes="Vector: pBRCDNAGFIAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_LIB=QH_EFGHU sunflower RHA280
 TAG_TISSUE=hulls
 TAG_SEQ=GCTAGTCGGG"

ASE COUNT 224 a 140 C 191 G 219 t 1 others
 RIGIN
 Alignment Scores:
 red. No.: 6.96e-132 Length: 775
 core: 1225.00 Matches: 228
 percent Similarity: 94.57% Conservative: 16
 est Local Similarity: 88.37% Mismatches: 14
 very Match: 46.85% Indels: 0
 B: 13 Gaps: 0

S-10-006-852-2 (1-502) x BU026594 (1-775)
 Y 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 b 1 GTTGGAACTGTGGCTCATCCGAAGCATCATGTTGGCCGGACTAGCTTCAAAAGAAA 60
 Y 141 TtpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 b 61 TGGCAAAACAAAATGAAAGCTCTTGGCAACCTTGGCAACCTTAACTGTAACCGGG 120
 Y 161 AlaAsnValGlnValCysTropGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 b 121 GCCAACGTTACAGGTTTGTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTGAAGAA 180
 Y 181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 Db 181 GTGAAGTTGAGGAAAGTTACTAGCTGATGATCCAGAAAGCTGTGGAGATGTTGAT 240
 Y 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 Db 241 GAAACACTAATTTGTGTGGTCTTCTGATCTGGTTCACCTCAATGTTGAATTTGAAGAT 300
 Y 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTropAspThrPro 240
 Db 301 GTTARGCGTTTAAATGACCTCTTGAATGAGAAAATGCCAAGACCCGATGGATACACCT 360
 Y 241 IleHisValAspAlaAspSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 Db 361 ATTCAATGGGCGTCCAGTGGAGGTTTATTGACCACTTATTACCCGGACTTGAA 420

Qy 261 TtpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 Db 421 TGGGATTTCCGGTTTCCATTGGTTAAGAGTATAAATCTTATGTTGTCACAAATATGGCTT 480
 Qy 281 ValTyrAlaGlyIleGlyTropValIleTropArgAsnLysGluAspLeuProGluGluLeu 300
 Db 481 GTTTACGCCGGAATTTGGTGGTTCATTTGGAGGACAAAGATGCTTGCCTGATGAATC 540
 Qy 301 IlePheHisIleAsnTyrLeuGlyValaAspGlnProThrPheThrLeuAsnPheSerLys 320
 Db 541 ATCTTTCACATCAACTATCTTGGTGTGATCAACCACTTTCACCTCAACTTCTCCAAA 600
 Qy 321 GlySerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 Db 601 GGTCTAGTCAATAAATGCTCAATCTATCAGTTCATTCGTTGGTTCGAGGGATAC 660
 Qy 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 Db 661 AAGAGTATCATGGAAAATGTCNCAAAAATGCAAAATGTTCTAAAGAAAGGTTGGAGAA 720
 Qy 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPhe 378
 Db 721 ACCGACGGTTCACATCATCTTAAGAACACCGCGTCCACACTTGTAGCTTC 774

RESULT 7
 CAS13481
 LOCUS
 DEFINITION
 CAB13481 786 bp mRNA linear EST 11-APR-2003
 CA48LN10IF-E8 Cabernet Sauvignon Leaf - CA48LN Vitis vinifera cDNA clone CA48LN10IF-E8 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 CAB13481.1 GI:26262418
 Vitis vinifera
 Vitis vinifera

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Doug Cook
 CABES Genome Facility
 UC Davis Department of Plant Pathology
 1 Shields Ave., Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drocock@ucdavis.edu
 Seq primer: GTTATCAGTCGACGGTACC.
 Location/Qualifiers
 1..786
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CA48LN10IF-E8"
 /sex="hermaphrodite"
 /dev_stage="Late season sample"
 /lab_hosts="DHSalpha"
 /clone_libs="Cabernet Sauvignon Leaf - CA48LN"
 /note="Organ: Leaf; Vector: pDNR; Site: 1: Sfil; Site 2:
 Sfil; CA48LN is a cDNA library of Cabernet Sauvignon
 leaves. The leaves were collected on September 20, 2001,
 in Napa Valley, California, and represent leaves in late
 season development. These leaves were asymptomatic and
 verified to be non-infected with the bacterial pathogen,
 Xylella fastidiosa, based on a diagnostic assay using PCR
 and Xylella-specific primer pairs. cDNAs were made by
 oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:

BQ971999
 BQ971999.1 GI:22389520
 EST.
 Helianthus annuus (common sunflower)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Helianthaceae; Helianthus.
 1 (bases 1 to 762)
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, W., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: QHB9 row: G column: 09.
 Location/Qualifiers
 source
 1..762
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA801"
 /db_xref="taxon:4232"
 /clone="QHB9G9"
 /lab_host="E.coli"
 /clone_lib="OH ABCDI sunflower RHA801"
 /note="Vector: pBRCDNASfiAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG LIB=OH ABCDI sunflower RHA801
 TAG TISSUE=hulls
 TAG_SEQ=GCTAGTCGGG"
 BASE COUNT 224 a 138 c 190 g 210 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.48e-129 Length: 762
 Score: 1203.00 Matches: 224
 Percent Similarity: 94.09% Conservative: 15
 Best Local Similarity: 88.18% Mismatches: 15
 Query Match: 46.01% Indels: 0
 DB: 13 Gaps: 0
 US-10-006-852-2 (1-502) x BQ971999 (1-762)
 Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 1 GTTGGAACTGTGGGCTCATCCGAAAGCCATCATGTGGCCGACTAGCTTTCAAAAGAAA 60
 Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 Db 61 TGGCAAAACAAATGAAAGACTCTTGGCAAAACCTTGGGACAAACCTAATTTGTATACCCGG 120
 Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 121 GCCAAGTTCAGTTTGTGGAGAAAATTTGCTCGGTATTTGAAGTGGAGTTGAAGGAA 180
 clone QHB9G09, mRNA sequence.

RESULF 9
 BQ971999
 QHB9G09.Y3.ab1 OH ABCDI sunflower RHA801 Helianthus annuus cDNA
 DEFINITION
 clone QHB9G09, mRNA sequence. EST 21-AUG-2002

```

181 VaLysLeuSerGluGlyTyrValMetAspProGlnAlaValAspMetValAsp 200
181 GTGAAGTGGAGGAAGGTTACTACGTGATCGATGATCGTGAAGAAGCTCTGGAGATGTTGAT 240
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
241 GAATAACTAATTTGTGGCTGCTACTTGGTTCACCTCCCAAGCCGATGGATTTGAAGAT 300
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrPro 240
301 GTTAAGCGTTTAAATGACCTCTTGAATGAGAAAATGCCAAGACCGGATGGGATACACCT 360
241 IleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
361 ATTCAATGGACCGTCAAGTGGAGGTTTATTGCACCAATTTATTACCCGGAACCTTGA 420
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
421 TGGATTTCCGGTTCCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480
281 ValTrpAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeu 300
481 GTTTACCGCGAATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
301 IlePheHisIleAsnTyrLeuGlyValAspGlnProThrPheThrLeuAsnPheSerLys 320
541 ATCTTTCCACATCAACTATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
601 GGTTCTAGTCAACAATGCTCAACTATCACTATCACTATCACTATCACTATCACTATCACT 660
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
661 AAGGACATCATGAAACTCTCAAGAAAATGCAAAATGCTTAAAGAAAGGTTTGGAGAAA 720
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValPro 374
721 ACCGGAGCGGTCAACATCACTCTTAAAGAACACCGGGCTCCCA 762

```

```

RESULT 10
E054156
OCUS
DEFINITION
  BE054156 803 bp mRNA linear EST 07-MAR-2001
  GA_Ea0034H22f Gossypium arboreum 7-10 dpa fiber library Gossypium
  arboreum cdNA clone GA_Ea0034H22f, mRNA sequence.
CCRSION
  BE054156.2 GI:13246539
EYWORDS
  EST.
SOURCE
  Gossypium arboreum
  Gossypium arboreum
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
  ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
  1 (bases 1 to 803)
AUTHORS
  Wang R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
  D., Wood, T.C., Leslie, A. and Wilkins, T.A.
TITLE
  An integrated analysis of the genetics, development, and evolution
  of the cotton fiber
JOURNAL
  Unpublished
COMMENT
  On Jun 8, 2000 this sequence version replaced gi:8381212.
  Contact: Wing RA
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Tel: 864 656 7288
  Fax: 864 656 4293
  Email: rwing@clemson.edu
  Seq primer: TAATACGACTCACTATAGG
  High quality sequence stop: 792.
  Location/Qualifiers
    source
      1..803
      /organism="Gossypium arboreum"
      /mol_type="mRNA"

```

```

/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0034H22f"
/tissue_type="Fibers isolated from bolls harvested 7-10
  dpa
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/notes="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 236 a 152 c 215 g 198 t 2 others
ORIGIN
Alignment Scores:
  1.76e-128 Length: 803
  1196.00 Matches: 220
  Score: 92.86% Conservative: 27
  Percent Similarity: 82.71% Mismatches: 19
  Best Local Similarity: 45.74% Indels: 0
  Query Match: 10 Gaps: 0
US-10-006-852-2 (1-502) x BE054156 (1-803)
QY 122 GlyThrValGlySerSerGluAlaIleWetLeuAlaGlyLeuAlaPheLysArgLysTyr 141
Db 2 GGGACAGTTGGTTCTCTGAGGCAATAATGCTGCGAGATTAGCCCTTAAAGGAAGTGG 61
QY 142 GlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAla 161
Db 62 CAACAAAAGATGAAATCACAGGAAAAACCGTATGATAAGCCCTAACATAGTCCCGGAGCG 121
QY 162 AsnValGlnValCysTyrGluLysPheAlaArgTyrPheGluValGluLeuLysGluVal 181
Db 122 AATGTCCAGGTTTGTGGAGAACGTTTCCGAGGATTTCCGAGGTTGAATTAAGGAAAGT 181
QY 182 LysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGlu 201
Db 182 AAGTTGAAGAGGATACTATGATGGACCTGTGAACCGGTTGAAATGGTTGACGAG 241
QY 202 AsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspVal 221
Db 242 AATACCATATGTTGCAGCAATCTCGGATCCACCCCTAACAGGGGAGTTTGAGAAATGG 301
QY 222 LysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrProIle 241
Db 302 AAGTCCCTTAATGAACTTCTCCAAAGAAAGAAATGAAGAAACCGGTTGGATPACCCCAATA 361
QY 242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTyr 261
Db 362 CATGTGGATGCTGTAGTGGAGGGTTTCATTTGCTCCCTTGTTTACCCGAAAATCGAATGG 421
QY 262 AspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeuVal 281
Db 422 GATTTCCCGCTCGCGTTAGTAAAAAAGCAATCAATGCAAGTGGGCAAGATATGGCCTTGT 481
QY 282 TyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeuIle 301
Db 482 TATGCTGTATCGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 541
QY 302 PheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLysGly 321
Db 542 TTTTCACTCAACTACCTTGGATCCGATCAGCCCAATTCACCTTANACTTCTCTAAAGGC 601
QY 322 SerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyrArg 341
Db 602 TCTAGTCAAAATCATAGTCAATAATTATCAGCTTATTCGGCTCGGTTTGGAGGATACAAG 661
QY 342 AsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlyLeuGluLysThr 361
Db 662 AGTATCATGGAAAATGCTATCGGAAATGCAAGAAATCTGAAGGAAAGGATAGAAAAACG 721
QY 362 GluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLys 381
Db 722 GGACGATTCGAAGTCGTTCCAAAGACGTTGGCGCTTCCCTTCGGGGCTTTTGCTCTAAA 781

```


Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozikeatgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: QHG7 row: H column: 16.

Location/Qualifiers
 1..786
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clone="QHGH16"
 /lab_host="E.coli"
 /clone_lib="OH_EFGHJ sunflower RHA280"
 /note="Vector: pBRCNDASfiAB; The library was constructed from 11 different sources of RNA from a single genotype. Incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG LIB=OH_EFGHJ sunflower RHA280
 TAG TISSUE=nulls
 TAG SEQ=CGTAGTCGGG"
 228 a 194 g 220 t 2 others

BASE COUNT 228 a 194 g 220 t 2 others
 RIGIN
 Alignment Scores:
 red. No.: 1.67e-127 Length: 786
 core: 1187.50 Matches: 227
 percent Similarity: 92.37% Conservatative: 15
 est Local Similarity: 86.64% Mismatches: 18
 query Match: 45.41% Indels: 2
 B: 13 Gaps: 1

S-10-006-852-2 (1-502) x BU027739 (1-786)
 Y 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 |||||
 b 1 GTTGGAACTGGGCTCACCAGGCATCATGTTGCCGGAGTAGCTTCAAAAGAAA 60
 Y 141 TrpGlnAsnLysArgLysAlaGluGlyProValAspLysProAsnIleValThrGly 160
 |||||
 b 61 TGGCAAAACAAATGAAAGCTCTGGCAACCTGGGACAACTAACATGTAAACGGG 120
 Y 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 |||||
 b 121 GCCAACGTTCCAGCTTGGGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180
 Y 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 |||||
 b 181 GTGAAGTTGAGGAAAGGTTACTACGTGATGGATCCCTGAGAAAGCTGTGGAGATGGTTGAT 240
 Y 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 |||||
 b 241 GAAACACTATTGTGGGCTGCTATCTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300
 Y 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 |||||
 b 301 GTTAAAGCGTTAAATGACCTCTTGAATGAGAAAATGCCAAGCCGGATGGATACACCT 360
 Y 241 IleHisValAspAlaAsnGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 |||||
 b 361 ATTCAATGGACGCTCAAGTGGAGGTTTATTGCCACATTTATTACCCGGAACTTGA 420
 Y 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 |||||
 b 421 TGGGATTCGGTTGCCATGGTTAAGATATAAATGTTAGTGGTCACAAAATATGGCCT 480

QY 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 |||||
 Db 481 GTTTACCGCCGAAATTTGGTGGTCAATTTGGAGGAACAAGATGACTTGCCTGATGAATC 540
 QY 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 |||||
 Db 541 APTCTTCACTCAACTAFTCTGGGCTGATCAACCACTTTCCTCACTTCAACTTCCAAA 600
 QY 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTy 340
 |||||
 Db 601 GGTTCAGTCAATAATTTGCTCAATCACTATCAGTTTCAGTTTCGTTGGGTTTCGAGGATA 660
 QY 340 rArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeu--G1 359
 |||||
 Db 661 CNAGAGCATCATGGAAACTCTCAAGAAAATGCAAAATGTTCTTAAAGAAAGGGTTTGGAG 720
 QY 359 ulysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSe 379
 |||||
 Db 721 AAACCCGGGACGGTTCACATCACTCTTAAAGACAACGGCTCCACTTGTAGCTTCTC 780
 QY 379 rLeu 380
 |||||
 Db 781 GTTA 784

RESULT 13
 BU025264
 LOCUS
 DEFINITION
 BU025264 clone QHF8J20, mRNA sequence.
 BU025264.1 GI:22460784
 EST.
 SOURCE
 Helianthus annuus (common sunflower)
 ORGANISM
 Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 REFERENCE
 1 (bases 1 to 746)
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavalle, D., Chevaller, P., Ziegler, J., Ellison
 P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
 Church, S., Jackson, B. and Bradford, K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 UNPUBLISHED
 CONTACT: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozikeatgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: QHF8 row: J column: 20.

Location/Qualifiers
 1..746
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clone="QHF8J20"
 /lab_host="E.coli"
 /clone_lib="OH_EFGHJ sunflower RHA280"
 /note="Vector: pBRCNDASfiAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library

construction can be obtained at <http://cgpdb.ucdavis.edu/>
 TAG_LIB=QH_EFGHJ sunflower RHA280
 TAG_TISSUE=hulls
 TAG_SEQ=GCTAGTCGGG"

BASE COUNT 219 a 131 c 185 g 211 t
 ALIGNMENT SCORES:
 Alignment Scores:
 Pred. No.: 3,01e-127 Length: 746
 Score: 1185.00 Matches: 220
 Percent Similarity: 95.16% Conservat: 16
 Best Local Similarity: 88.71% Mismatches: 12
 Query Match: 45.32% Indels: 0
 Gaps: 13

US-10-006-852-2 (1-502) x BU025264 (1-746)
 Y 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 b 1 GTTGGAACTGTGGGCTCATCCGAAGCCATCANGTTGGCCGGACTAGCTTCAAAAGAAA 60
 Y 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 b 61 TGGCAAAACAATGAAAGACTCTTGGCAAACTTGGACAACTTAACATTGTAAACCGG 120
 Y 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 b 121 GCCAAGCTTCAGCTTGTGGGAGAAATTTCTCGGTAATTTTGAAGTGGAGTTGAAGAA 180
 Y 181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 b 181 GTGAAGTTGAGGAAAGGTTACTACGTAGTGGATCTTGGAGAAAGCTGTGGAGATGGTTGAT 240
 Y 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 b 241 GAAAACACTATTTGTGGCTGCTAICTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300
 Y 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspPhePro 240
 b 301 GTTAAAGCTTTAAATGACCTTCTGATTGAGAAAATGCCAAGACCAGGATGGGATACACCT 360
 Y 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 b 361 ATTCAATGTGACGCTCAAGTGGAGGTTTATTGCCACCAATTAATACCAGGAACTTGA 420
 Y 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 b 421 TGGATTTCCGGTTGCCATTGGTTAGAGATATAAATGTTAGTGTACAAATATGGGCTT 480
 Y 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 b 481 GTTTACCGGAAATGGGTGGGTCATTTGGAGGAAACAAGATGACITTCCTGTATGAATC 540
 Y 301 IlePheHisIleAsnTyrLeuGlyValAspGlnProThrPheThrLeuAsnPheSerLys 320
 bb 541 ATCTTTCACANCACTAATCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 Y 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 bb 601 GGTTCCTAGTCAAAATTAATGCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 660
 Y 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 bb 661 AAGACATCTGAAAACTGTCTGAAAAATGCAAAATGTTCTAAAGAAAGGTTTGGAAA 720
 Y 361 ThrGluArgPheAsnIleValSer 368
 bb 721 ACCGACGGTTCAACATCACTCT 744

RESULT 14
 BU027680
 LOCUS
 DEFINITION QHG7B13.y9.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
 746 bp mRNA linear EST 23-AUG-2002

clone QHG7B13, mRNA sequence.
 BU027680
 BU027680.1 GI:22463200
 EST.
 SOURCE Helianthus annuus (common sunflower)
 ORGANISM Helianthus annuus

REFERENCE 1 (bases 1 to 746)
 AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
 Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
 ,P., Kolkman,J., Siabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
 Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished

TITLE Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig2686, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: QHG7 row: B column: 13.

FEATURES
 Location/Qualifiers
 source
 1..746
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clone="QH_G7B13"
 /lab_host="E.coli"
 /clone_lib="QH_EFGHJ sunflower RHA280"
 /note="Vector: pBERDNASFIAB; The library was constructed
 from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at <http://cgpdb.ucdavis.edu/>
 TAG_TISSUE=hulls
 TAG_SEQ=GCTAGTCGGG"

BASE COUNT 219 a 131 c 185 g 211 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,01e-127 Length: 746
 Score: 1185.00 Matches: 220
 Percent Similarity: 95.16% Conservat: 16
 Best Local Similarity: 88.71% Mismatches: 12
 Query Match: 45.32% Indels: 0
 Gaps: 13

US-10-006-852-2 (1-502) x BU027680 (1-746)

QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 1 GTTGGAACTGTGGGCTCATCCGAAGCCATCANGTTGGCCGGACTAGCTTCAAAAGAAA 60
 QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 Db 61 TGGCAAAACAATGAAAGACTCTTGGCAAACTTGGACAACTTAACATTGTAAACCGG 120

QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 121 GCCAAGCTTCAGCTTGTGGGAGAAATTTCTCGGTAATTTTGAAGTGGAGTTGAAGAA 180

FEATURES source

Location/Qualifiers

1. .746

/organism="Helianthus annuus"

/mol_type="mRNA"

/cultivar="RHA280"

/db_xref="taxon:4232"

/clone="QHGI0J06"

/lab_host="E.coli"

/clone_lib="QH_EFGHJ sunflower RHA280"

/note="vector: pBRCNDNAS1AB; The library was constructed from 11 different sources of RNA from a single genotype. separate cDNAs were generated using primers that incorporated unique 5', and 3', tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/TAG_LIB-OH_EFGHJ_sunflower_RHA280

TAG_LIB-OH_EFGHJ_sunflower_RHA280

TAG_TISSUE=nulls

TAG_SEQ=GCTAGTGGGG

BASE COUNT 218 a 132 c 185 g 211 t

ORIGIN

Alignment Scores: 3.94e-127 Length: 746

Pred. No.: 1184.00 Matches: 220

Score: 95.16% Conservative: 16

Best Local Similarity: 88.71% Mismatches: 12

Query Match: 45.28% Indels: 0

DB: 13 Gaps: 0

US-10-006-852-2 (1-502) x BU025649 (1-746)

QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140

Db 1 GTTGAACATGTGGGCTCATCCGAGCCATCATGTGGCCGGACTAGCTTCAAAAAGAAA 60

QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160

Db 61 TGGCAAAAACAAAATGAAAAGCTCTTGGCAAACTTGGGACAAAACCTAACATTTGTAACCCGG 120

QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180

Db 121 GCCAACGTTTCAGTTTGTGGGAGAAAATTCCTCGTATTTTGAAGTGGAGTTGAAGAA 180

QY 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200

Db 181 GTGAAGTTGAGGGAAGTTACTACGCTGATGGATCCTGAGAAAAGCTGTGGAGATGTTGAT 240

QY 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220

Db 241 GAAAAACATATTGTGTGGCTGCTAATCCCTCCACCCCAATGGTGAATTTGAAAGAT 300

QY 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240

Db 301 GTTAAAGCGTTTAAATGACCTCTTGTATGAGAAAATGCCAAGACCCGATGGATACACT 360

QY 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaPropheLeuTyrProGluLeuGlu 260

Db 361 ATTCAATGTGGAGCTGCAAGTGGAGGTTTATTGACCACATTTATTACCCGGAACTTCAA 420

QY 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280

Db 421 TGGGATTTCCGTTCCCATTTGGTTAAGAGATATAAATGTAGTGTGCACAAAATATGGGCTT 480

QY 281 ValTyrAlaGlyIleGlyTrpValIleLeuTrpArgAsnLysGluAspLeuProGluLeu 300

Db 481 GTTACCCCGGAATTTGGTGGGTCAITTTGGAGAAAACAAAGATGACTTCCCTGATGAATC 540

QY 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320

Db 541 ATCTTTCACATCAACTATCTTGGTGTGATCAACCAACTTCACTCAACTTCTCCAAA 720

QY 361 ThrGluArgPheAsnIleValSer 368

Db 721 ACCGGACGGTTCAACATCATCTCT 744

181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200

181 GTGAAGTTGAGGGAAGTTACTACGCTGATGGATCCTGAGAAAAGCTGTGGAGATGTTGAT 240

201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220

241 GAAAAACATATTGTGTGGCTGCTAATCCCTCCACCCCAATGGTGAATTTGAAAGAT 300

221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240

301 GTTAAAGCGTTTAAATGACCTCTTGTATGAGAAAATGCCAAGACCCGATGGATACACT 360

241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaPropheLeuTyrProGluLeuGlu 260

361 ATTCAATGTGGAGCTGCAAGTGGAGGTTTATTGACCACATTTATTACCCGGAACTTCAA 420

261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280

421 TGGGATTTCCGTTCCCATTTGGTTAAGAGATATAAATGTAGTGTGCACAAAATATGGGCTT 480

281 ValTyrAlaGlyIleGlyTrpValIleLeuTrpArgAsnLysGluAspLeuProGluLeu 300

481 GTTACCCCGGAATTTGGTGGGTCAITTTGGAGAAAACAAAGATGACTTCCCTGATGAATC 540

301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320

541 ATCTTTCACATCAACTATCTTGGTGTGATCAACCAACTTCACTCAACTTCTCCAAA 600

321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340

601 GGTTCAGTCAAAATATGCTCAATCACTATCACTTCACTTCCCTGGTTCGAGGGANAC 660

341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360

661 AAGAGCATCATGAAACTGTCAAGAAAATGCAAAATGTTCTAAAGAAAAGTTGGAGAAA 720

361 ThrGluArgPheAsnIleValSer 368

721 ACCGGACGGTTCAACATCATCTCT 744

BU025649 746 bp mRNA linear EST 23-AUG-2002

QHGI0J06.yg.ab1 OH EFGHJ sunflower RHA280 Helianthus annuus cDNA

clone QHGI0J06, mRNA sequence.

BU025649

BU025649.1 GI:22461169

Helianthus annuus (common sunflower)

Helianthus annuus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteraceae; Asteroidae; Heliantheae; Helianthus.

1 (bases 1 to 746)

Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.

Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compgenomics.ucdavis.edu/

Unpublished

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmurdson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]

belongs to contig QH_CA_Contig2686, see http://cgdb.ucdavis.edu/

for details.

Plate: QHGI0 row: J column: 06.

```

Y       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
      ||| ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
b       601 GGTTC TAGTCAAAATAATTGCTCAATACTATCAGTTCATTGGCTTGGGTTTCGAGGATAC 660
      ||| ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValIleuArgGluGlyLeuGluLys 360
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
b       661 AAGAGC ATCATGGAAACTGTCAGAAATGCCAAATGTCCTACAGAAAGTTTGAGAAA 720
      ||| ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y       361 ThrGluArgPheAsnIleValSer 368
      ||| ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b       721 ACCGGACGGTTCACATCCTCT 744
      ||| ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: October 22, 2003, 17:27:39
 Job time : 2452 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Protein - protein search, using sw model

run on: October 22, 2003, 14:50:13 ; Search time 69 Seconds
(without alignments)
1877.426 Million cell updates/sec

title: US-10-006-852-2

effect score: 2615
sequence: 1 MWLSHAVSESDVSVHSTFAS.....DIITGKMKFVADRKKTGIC 502

coring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

searched: 830525 seqs, 258052604 residues

total number of hits satisfying chosen parameters: 830525

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriopl.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Match	Length	DB	ID	Description
1	2240	85.7	496	10	O81102	O81102 nicotiana t
2	2233	85.4	496	10	O81KR4	O81KR4 nicotiana t
3	2232	85.4	496	10	P93369	P93369 nicotiana t
4	2228	85.2	496	10	O9AT17	O9AT17 nicotiana t
5	2199	84.1	496	10	O81101	O81101 nicotiana t
6	2191.5	83.8	513	10	O8LLP2	O8LLP2 cryza sativ
7	2186.5	83.6	491	10	O94KK8	O94KK8 nicotiana t
8	2135.5	81.7	493	10	O9ZPS3	O9ZPS3 arabidopsis
9	2126	81.3	494	10	O944L6	O944L6 arabidopsis
10	2080	79.5	500	10	O9ZFS4	O9ZFS4 arabidopsis
11	2056.5	78.6	501	10	O9AQU4	O9AQU4 cryza sativ
12	1969	75.3	494	10	O9LSH2	O9LSH2 arabidopsis
13	1965	75.1	494	10	O8LFR4	O8LFR4 arabidopsis
14	1845.5	70.6	419	10	O8RXH0	O8RXH0 arabidopsis
15	1844	70.5	500	10	O9AR41	O9AR41 cryza sativ
16	1321	50.5	520	3	Q8X0B0	Q8X0B0 neurospora

ID	Accession	Score	Match	Length	Description
17	1318	50.4	514	3	Q9UVD7
18	1251.5	47.9	467	16	P73043
19	1110	42.4	460	16	O06249
20	1108	42.4	475	16	O9X8J5
21	1100	42.1	468	17	O8TPG4
22	1070.5	40.3	282	10	Q8RV54
23	1008	38.5	489	16	Q8FHG5
24	1007.5	38.5	464	16	Q8XIQ6
25	952	36.4	464	2	Q8QF15
26	635.5	24.3	304	16	O8YBJ0
27	613.5	23.5	177	10	O9LSH6
28	487	18.6	219	2	Q8L3L1
29	401	15.3	163	2	Q8L3L0
30	379	14.5	157	2	Q8LC52
31	335.5	12.8	167	16	O8YBI9
32	325	12.4	589	3	Q05567
33	317	12.1	372	17	O8TV92
34	317	12.1	576	3	Q8X074
35	314	12.0	384	17	Q9UZD5
36	312	11.9	363	17	O27188
37	311	11.9	367	17	O28275
38	305	11.7	383	17	O58679
39	299.5	11.5	545	5	O9V7Y2
40	290.5	11.1	395	17	Q8TUQ9
41	288	11.0	552	5	Q9Y194
42	271	10.4	398	17	Q8PXA5
43	270	10.3	544	10	O9C509
44	269.5	10.3	488	17	O28946
45	264.5	10.1	542	5	Q17456

ALIGNMENTS

RESULT 1

O81102 PRELIMINARY; PRT; 496 AA.
 AC O81102;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Glutamate decarboxylase isozyme 1.
 GN NRGAD1.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]_TaxID=4097;
 RP SEQUENCE FROM N.A.
 RA Yun S.J., Oh S.H.;
 RT "Cloning and characterization of a tobacco cDNA encoding
 calcium/calmodulin-dependent glutamate decarboxylase";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -|- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
 DR EMBL; AF020425; AAC24195.1; .
 DR InterPro: IPR002129; Pyridoxal dec.
 DR Pfam: PF00282; Pyridoxal dec; I.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 SQ SEQUENCE 496 AA; 55963 MW; 4E8492F5DF0BDA8A CRC64;

Query Match 85.7%; Score 2240; DB 10; Length 496;
 Best Local Similarity 85.1%; Pred. No. 4, 1e-161;
 Matches 428; Conservative 35; Mismatches 32; Indels 8; Gaps 3;
 QY 1 MWLSHAVSESDVSVHSTFASRYVRSLSFRKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
 Db 1 MWLSKTSSESDVSIHSTFASRYVRSLSFRKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
 QY 61 ASFVITWPEPCDKLIMSSINKNYVDMDEYPTTTELQRCNVMIAHLNAPLEAEATVG 120

61 ASFVTTWMEPECKLMMS INKNYVDMDEYVPTTELQNCVNMIAHLFNAPLGGGTAVG 120
121 VGTVGSSEAIMLAGLAFKRWKQKKAEGKFPVDPKPNIVTGANVQVCWEKFPARYFEVELKE 180
121 VGTVGSSEAIMLAGLAFKRWKQKKAEGKFPVDPKPNIVTGANVQVCWEKFPARYFEVELKE 180
121 VGTVGSSEAIMLAGLAFKRWKQKKAEGKFPVDPKPNIVTGANVQVCWEKFPARYFEVELKE 180
181 VKLSGGYVMPDPOQAVDMVDENTICVADIIIGSTLNGEFDVKKLLNDLLVEKNKKTGWDTP 240
181 VKLSGGYVMPDPOQAVDMVDENTICVADIIIGSTLNGEFDVKKLLNDLLVEKNKKTGWDTP 240
241 IHVDAASGGFTAPFLYPELWDFRLPLVKSINVSGHKYGLVYAGIHWAIWRNKEDLPEEL 300
241 IHVDAASGGFTAPFLYPELWDFRLPLVKSINVSGHKYGLVYAGIHWAIWRNKEDLPEEL 300
301 IFHINYLGAQDPTFTLNFSGSSQVIAQYQLIRLGHGEGYRNWENCRENMIVRLEGLEK 360
301 IFHINYLGAQDPTFTLNFSGSSQVIAQYQLIRLGHGEGYRNWENCRENMIVRLEGLEK 360
361 TERFNIVSKDEGVLVAFSLKDSCHTEFEISDMRLRYGWIIVPAYTMPNAQHIITVLRV 420
361 SGRFNLSKEIGVPLVAFSLKDSCHTEFEISDMRLRYGWIIVPAYTMPNAQHIITVLRV 420
421 IREDFSRTLAERLVIDIEKVMREDELPSRVHKSILGQEKSESNSDNLMTVTKSDIDK 480
421 IREDFSRTLAERLVIDIEKVMREDELPSRVHKSILGQEKSESNSDNLMTVTKSDIDK 480
481 QRDIIITGKKFVAD-RKKTSGIC 502
474 QLEITTAWKKFVADKKKKTNGVC 496

RESULT 2
8LKR4 PRELIMINARY; PRT; 496 AA.

D QBLKR4
C QBLKR4
T 01-OCT-2002 (TReMBLrel. 22, Created)
T 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
T 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
E Glutamate decarboxylase.
S Nicotiana tabacum (Common tobacco).
C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
C Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
X NCBI_TaxID=4097;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=cv. Samsun NN;
A McLean M.D., Yevtushenko D., Deschene A., Van Cauwenberghs O.R.,
A Makhmoudova A., Potter J.W., Bown A.W., Shelp B.J.;
I "Transgenic tobacco plants overexpressing glutamate decarboxylase are
I resistant to the root-knot nematode."
L Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
C -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C TYRDC).
R EMBL; AF506366; AA48129.1; -
R InterPro; IPR002129; Pyridoxal_dec.
R Pfam; PF00282; pyridoxal_dec; I.
W Decarboxylase; Lyase; Pyridoxal phosphate.
Q SEQUENCE 496 AA; 55966 MW; 4771CF71BCA478349 CRC64;

Query Match 85.4%; Score 2233; DB 10; Length 496;
Best Local Similarity 84.9%; Pred. No. 1.4e-160;
Matches 427; Conservative 35; Mismatches 33; Indels 8; Gaps 3;

Y 1 MVLSHAVSESDVSVHSTFASRYVTSIPLRFKMPNSIPKEAAAYQIINDELMLDGNPRLNL 60
b 1 MVLSKTASESDVSVHSTFASRYVTSIPLRFKMPNSIPKEAAAYQIINDELMLDGNPRLNL 60
Y 61 ASFVTTWMEPECKLIMSS INKNYVDMDEYVPTTELQNCVNMIAHLFNAPLBBEATAVG 120
b 61 ASFVTTWMEPECKLIMSS INKNYVDMDEYVPTTELQNCVNMIAHLFNAPLGGGTAVG 120

121 VGTVGSSEAIMLAGLAFKRWKQKKAEGKFPVDPKPNIVTGANVQVCWEKFPARYFEVELKE 180
121 VGTVGSSEAIMLAGLAFKRWKQKKAEGKFPVDPKPNIVTGANVQVCWEKFPARYFEVELKE 180
181 VKLSGGYVMPDPOQAVDMVDENTICVADIIIGSTLNGEFDVKKLLNDLLVEKNKKTGWDTP 240
181 VKLSGGYVMPDPOQAVDMVDENTICVADIIIGSTLNGEFDVKKLLNDLLVEKNKKTGWDTP 240
241 IHVDAASGGFTAPFLYPELWDFRLPLVKSINVSGHKYGLVYAGIHWAIWRNKEDLPEEL 300
241 IHVDAASGGFTAPFLYPELWDFRLPLVKSINVSGHKYGLVYAGIHWAIWRNKEDLPEEL 300
301 IFHINYLGAQDPTFTLNFSGSSQVIAQYQLIRLGHGEGYRNWENCRENMIVRLEGLEK 360
301 IFHINYLGAQDPTFTLNFSGSSQVIAQYQLIRLGHGEGYRNWENCRENMIVRLEGLEK 360
361 TERFNIVSKDEGVLVAFSLKDSCHTEFEISDMRLRYGWIIVPAYTMPNAQHIITVLRV 420
361 SGRFNLSKEIGVPLVAFSLKDSCHTEFEISDMRLRYGWIIVPAYTMPNAQHIITVLRV 420
421 IREDFSRTLAERLVIDIEKVMREDELPSRVHKSILGQEKSESNSDNLMTVTKSDIDK 480
421 IREDFSRTLAERLVIDIEKVMREDELPSRVHKSILGQEKSESNSDNLMTVTKSDIDK 480
481 QRDIIITGKKFVAD-RKKTSGIC 502
474 QLEITTAWKKFVADKKKKTNGVC 496

RESULT 3
P93369 PRELIMINARY; PRT; 496 AA.

AC P93369
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Glutamate decarboxylase.
GN NIGADI.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=xanthi;
RA Dharmasiri M.A.N., Lu Y.T., Harrington H.M.;
RT "Cloning and sequencing of a tobacco cDNA encoding glutamate
RT decarboxylase."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
DR EMBL; U54774; AA40608.1; -
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; I.
KW Decarboxylase; Lyase; Pyridoxal phosphate.
SQ SEQUENCE 496 AA; 56035 MW; C48492F4BF0BDABA CRC64;

Query Match 85.4%; Score 2232; DB 10; Length 496;
Best Local Similarity 84.9%; Pred. No. 1.6e-160;
Matches 427; Conservative 35; Mismatches 33; Indels 8; Gaps 3;

QY 1 MVLSHAVSESDVSVHSTFASRYVTSIPLRFKMPNSIPKEAAAYQIINDELMLDGNPRLNL 60
Db 1 MVLSKTASESDVSVHSTFASRYVTSIPLRFKMPNSIPKEAAAYQIINDELMLDGNPRLNL 60
QY 61 ASFVTTWMEPECKLIMSS INKNYVDMDEYVPTTELQNCVNMIAHLFNAPLBBEATAVG 120
Db 61 ASFVTTWMEPECKLIMSS INKNYVDMDEYVPTTELQNCVNMIAHLFNAPLGGGTAVG 120
QY 121 VGTVGSSEAIMLAGLAFKRWKQKKAEGKFPVDPKPNIVTGANVQVCWEKFPARYFEVELKE 180

121 VGTVGSSEALMAGLAFKRWQNKKAQKPCPKPNIVTGANVQVCWEKFAFYFEVELKE 180
 181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVKNKKTGWDTTP 240
 181 VKLSDGYVMDPEKAVEMVDENTICVAAAILGSTLNGEFEDVKRLNDLLIEKNKKTGWDTTP 240
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPDEL 300
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPDEL 300
 301 IFHINVLGADQPTFTLNFSGSSQVIAQYQLIRLHGEGYRNVMENRENMIVLREGLEK 360
 301 IFHINVLGADQPTFTLNFSGSSQVIAQYQLIRLHGEGYRNVMENRENMIVLREGLEK 360
 361 TERFNIVSKDGVPLVAFSLKDSCHTEFEISDMLRRYGWIPAYTMPNACHITVLRVY 420
 361 SGRFNIIISKEIGVPLVAFSLKDSCHTEFEISDMLRRYGWIPAYTMPNACHITVLRVY 420
 421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLQEKSESDNLMVTVKSSDIDK 480
 421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLQEKSESDNLMVTVKSSDIDK 480
 481 QRDITGWKKFVAD-RKKTSGIC 502
 474 QLEITAAWKKFVADKKKKTNGVC 496

RESULT 4
 9AT17 PRELIMINARY; PRT; 496 AA.
 O9AT17
 O9AT17
 01-JUN-2001 (TREMBlrel. 17, Created)
 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 Glutamate decarboxylase isozyme 3 (EC 4.1.1.15).
 Nicotiana tabacum (Common tobacco).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=cv. Samsun NN;
 Yevtushenko D., McJann M.D., Peiris S.B., Van Cauwenbergh O.R.,
 Shelp B.J.;
 "Two isoforms of tobacco glutamate decarboxylase are regulated by
 calcium/calmodulin and differ in organ distribution."
 Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 C -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 C -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 TYRDC).
 C EMBL; AF352732; AAK18620.1; -;
 C InterPro; IPR002129; Pyridoxal_dec.
 C Pfam; PF00282; pyridoxal_dec; 1.
 C Decarboxylase; Lyase; PYRIDOXAL phosphate.
 Q SEQUENCE 496 AA; 55875 MW; C61BCD3531827BB CRC64;

Query Match 85.2%; Score 2228; DB 10; Length 496;
 Best Local Similarity 84.7%; Pred. No. 3.3e-160;
 Matches 426; Conservative 35; Mismatches 34; Indels 8; Gaps 3;

Y 1 MVLSHAVSESDVSHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLML 60
 bb 1 MVLKSTASESDVSHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLML 60
 Y 61 ASFVTTWMEPECKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEATVG 120
 bb 61 ASFVTTWMEPECKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEATVG 120
 Y 121 VGTVGSSEALMAGLAFKRWQNKKAQKPCPKPNIVTGANVQVCWEKFAFYFEVELKE 180
 bb 121 VGTVGSSEALMAGLAFKRWQNKKAQKPCPKPNIVTGANVQVCWEKFAFYFEVELKE 180

181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVKNKKTGWDTTP 240
 181 VKLSDGYVMDPEKAVEMVDENTICVAAAILGSTLNGEFEDVKRLNDLLIEKNKKTGWDTTP 240
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPDEL 300
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPDEL 300
 301 IFHINVLGADQPTFTLNFSGSSQVIAQYQLIRLHGEGYRNVMENRENMIVLREGLEK 360
 301 IFHINVLGADQPTFTLNFSGSSQVIAQYQLIRLHGEGYRNVMENRENMIVLREGLEK 360
 361 TERFNIVSKDGVPLVAFSLKDSCHTEFEISDMLRRYGWIPAYTMPNACHITVLRVY 420
 361 SGRFNIIISKEIGVPLVAFSLKDSCHTEFEISDMLRRYGWIPAYTMPNACHITVLRVY 420
 421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLQEKSESDNLMVTVKSSDIDK 480
 421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLQEKSESDNLMVTVKSSDIDK 480
 481 QRDITGWKKFVAD-RKKTSGIC 502
 474 QLEITAAWKKFVADKKKKTNGVC 496

RESULT 5
 081101 PRELIMINARY; PRT; 496 AA.
 ID 081101;
 AC 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glutamate decarboxylase isozyme 2.
 GN NTGAD2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98302498; PubMed=9638642;
 RA Yun S.J., Oh S.H.;
 RT "Cloning and characterization of a tobacco cDNA encoding
 calcium/calmodulin-dependent glutamate decarboxylase."
 RL Mol. Cells 8:125-129(1998).
 CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 TYRDC).
 CC EMBL; AF020424; AAC39483.1; -;
 DR InterPro; IPR002129; Pyridoxal_dec.
 DR Pfam; PF00282; pyridoxal_dec; 1.
 DR Decarboxylase; Lyase; PYRIDOXAL phosphate.
 Q SEQUENCE 496 AA; 55931 MW; 5B9D2C8E12560D27 CRC64;

Query Match 84.1%; Score 2199; DB 10; Length 496;
 Best Local Similarity 84.1%; Pred. No. 5.2e-158;
 Matches 423; Conservative 34; Mismatches 38; Indels 8; Gaps 3;

Y 1 MVLSHAVSESDVSHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLML 60
 Db 1 MVLKSTASESDVSHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLML 60
 QY 61 ASFVTTWMEPECKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEATVG 120
 Db 61 ASFVTTWMEPECKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEATVG 120
 QY 121 VGTVGSSEALMAGLAFKRWQNKKAQKPCPKPNIVTGANVQVCWEKFAFYFEVELKE 180
 Db 121 VGTVGSSEALMAGLAFKRWQNKKAQKPCPKPNIVTGANVQVCWEKFAFYFEVELKE 180
 QY 181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVKNKKTGWDTTP 240

181 VKLSDGYVMDPKAVEMVDENTICVAAALGSLTNGEFEDVVKRLNDLLIEKKNKGTGMDTP 240
 241 IHVDAASGGFIAPFLPELEWDFRPLVKSINVSCHKYGLVAVAGIHWIWRNKEDLPBEL 300
 241 IHVDAASGGFIAPFLPELEWDFRPLVKSINVSCHKYGLVAVAGIHWIWRNKEDLPBEL 300
 301 IFHINYLGDQPTFTLNFSGSSQVIAQYQIRLIRLHGHGFRVNMENCRNMIIVLREGLEK 360
 301 IFHINYLGDQPTFTLNFSGSSQVIAQYQIRLIRLHGHGFRVNMENCRNMIIVLREGLEK 360
 361 TERFNIVKDEGVLVAFSLKSSCHTEREISDMLRRYGWIYPATMPNAOHIIVLVRVV 420
 361 SGFNFIISKEIGVPLVAFSLKSSCHTEREISDMLRRYGWIYPATMPNAOHIIVLVRVV 420
 421 IREDFSRITLAERLVIDIEKVMRELDLSPRVIHKISLGOEKSESNDNIMLVVKKSDIDK 480
 421 IREDFSRITLAERLVIDIEKVMRELDLSPRVIHKISLGOEKSESNDNIMLVVKKSDIDK 480
 481 QRDIIITGKWKVAD-RKKTSGIC 502
 474 QLEITTAWLKFKVADKKKTNGVC 496

RESULT 6
 SLLP2 PRELIMINARY; PRF; 513 AA.
 D Q8LLP2
 C Q8LLP2;
 I 01-OCT-2002 (TrEMBLrel. 22, Created)
 T 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 I 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 E Putative glutamate carboxylase.
 S OSJNB0031009.06.
 N Oryza sativa (japonica cultivar-group).
 C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 C Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 X Ehrhartoideae; Oryzaceae; Oryza.
 X NCBI_TaxID=39947;
 N [1]
 N SEQUENCE FROM N.A.
 P Eastman A.P., Smith S.C., Bertin N., Liang C., Gingle A.R.,
 A Pratt L.H., Cordonnier-Pratt M.-M.;
 A "Sequencing of clone OSJNB0031009."
 T Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 L [2]
 P SEQUENCE FROM N.A.
 P Eastman A.P., Smith S.C., Bertin N., Liang C., Najjar F.Z., Pratt L.H.,
 A Cordonnier-Pratt M.-M.;
 A "Untitled."
 T Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 L C
 C -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
 C EMBL; AF377946; AAM47304.1; -;
 R Gramine; Q8LLP2; -;
 R InterPro; IPR002129; Pyridoxal dec.
 R Pfam; PF00282; Pyridoxal dec; 1.
 W Decarboxylase; Lyase; Pyridoxal phosphate.
 Q SEQUENCE 513 AA; 57378 MW; 131DFD62F6R0CE87 CRC64;

Query Match 83.8%; Score 2191.5; DB 10; Length 513;
 Best Local Similarity 79.7%; Pred. No. 2e-157;
 Matches 417; Conservative 38; Mismatches 37; Indels 31; Gaps 2;

1 MVLHAVSESDVSVHSTFASRYVRTSLPR-----FKMPENSIEPK 39
 1 MVLHGVSGSDES VHSSTFASRYVRTSLPRHARGLSRAPLAPIDSVIDWEFRMEQSIK 60
 40 EAAVQIINDELMIDGNPRNLASVVTWMEPCDKLIMSSINKNYVMDDEYPTVTELQNR 99
 61 EAAVQIINDELMIDGNPRNLASVVTWMEPCDKLIMSSINKNYVMDDEYPTVTELQNR 120
 100 CVNMIHLFNAPLEAEATYGVGTGSSSEAIMLAGLAFKRWQNKKAEGKPKVDKPNIVT 159

121 CVNMIHLFNAPLEAEATYGVGTGSSSEAIMLAGLAFKRWQNKKAEGKPKVDKPNIVT 180
 160 GANVOVCWEKFAFYFVELKEVLSGEGYVMDPOQAVDMVDENTICVADILGSLTNGEFE 219
 181 GANVOVCWEKFAFYFVELKEVLSGEGYVMDPOQAVDMVDENTICVADILGSLTNGEFE 240
 220 DVKLLNDLLVEKNETGWDTPIHVDAASGGFIAPFLPELEWDFRPLVKSINVSCHKY 279
 241 DVKLLNDLLTKNAETGWDTPIHVDAASGGFIAPFLPELEWDFRPLVKSINVSCHKY 300
 280 LVAAGIHWIWRNKEDLPBELIHNINVLGDQPTFTLNFSGSSQVIAQYQIRLIRLHGHG 339
 301 LVAAGIHWIWRNKEDLPBELIHNINVLGDQPTFTLNFSGSSQVIAQYQIRLIRLHGHG 360
 340 YRNVMENCRNMIIVLREGLEKTERFNIVSKDEGVPLVAFSLKSSCHTEREISDMLRRYG 399
 361 YKNIMENCRNMIIVLREGLEKTERFNIVSKDEGVPLVAFSLKSSCHTEREISDMLRRYG 420
 400 WIVPAYMPPNAOHIIVLVRVIREDFSRITLAERLVIDIEKVMRELDLSPRVIHKISLGO 459
 421 WIVPAYMPPNAOHIIVLVRVIREDFSRITLAERLVIDIEKVMRELDLSPRVIHKISLGO 473
 460 EKSESNDNIMLVVKKSDIDKORDIITGKWKVADKKKTSGIC 502
 474 ---ANGGDAASASEREMEKQREVISLWKAVALKKTNGVC 513

RESULT 7
 O94KK8 PRELIMINARY; PRF; 491 AA.
 AC O94KK8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Glutamate decarboxylase isozyme 4 (EC 4.1.1.15).
 OS Nicotiana tabacum (Common tobacco).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yevrushenko D., McLean M.D., Peiris S.E., Van Cauwenbergh O.R.,
 RA Shelp B.J.;
 RT "Two isoforms of tobacco glutamate decarboxylase are regulated by calcium/calmodulin and differ in organ distribution."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
 DR EMBL; AF353615; AAK38667.1; -;
 DR InterPro; IPR002129; Pyridoxal_dec.
 DR Pfam; PF00282; Pyridoxal dec; 1.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 SQ SEQUENCE 491 AA; 55950 MW; 51CB78B3EF3F4E35 CRC64;

Query Match 83.6%; Score 2186.5; DB 10; Length 491;
 Best Local Similarity 84.9%; Pred. No. 4.5e-157;
 Matches 421; Conservative 27; Mismatches 39; Indels 9; Gaps 1;

1 MVLHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKGAAYQIINDELMIDGNPRNL 60
 1 MVLKTSSES DVS VHSSTFASRYVRTSLPRFMAENSIPKAAAFQIINDELMIDGNPRNL 60
 61 ASFVTTWMEPCDKLIMSSINKNYVMDDEYPTVTELQNR CVNMIHLFNAPLEAEATV 120
 61 ASFVTTWMEPCDKLIMSSINKNYVMDDEYPTVTELQNR CVNMIHLFNAPLEAEATV 120
 121 VGTGSSSEAIMLAGLAFKRWQNKKAEGKPKVDKPNIVTGANVOVCWEKFAFYFVELKE 180
 121 VGTGSSSEAIMLAGLAFKRWQNKKAEGKPKVDKPNIVTGANVOVCWEKFAFYFVELKE 180
 181 VKLSEGYVMDPKAVEMVDENTICVADILGSLTNGEFEDVVKRLNDLLIEKKNKGTGMDTP 240

181 VKLREGYVMDPVOAEMVDENTICVAAILGSLTNGEFEDVKLLNDLLLEKKNKQTGMNTP 240
 241 IHVDAASGGFIAFLPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRWKNKEDLPEEL 300
 241 IHVDAASGGFIAFLPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRWIKQDLPPEEL 300
 301 IFHINVLGADOPTFTLNFSGSSQVIAQYQYLIQLRGLGHEGFRVNMENCRNMIIVLREGLEK 360
 301 IFHINVLGADOPTFTLNFSGSSQVIAQYQYLIQLRGLGHEGFRVNMENCRNMIIVLREGLEK 360
 361 TERFNIVSKDEGVPLVAFSLKDSCHTEFEISDMRLRYGWIVPAYTWPNAQHITVLRVY 420
 361 TGRFNIVSKDEGVPLVAFSLKDSCHTEFEISDMRLRYGWIVPAYTWPNAQHITVLRVY 420
 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNDMLMTVTKKSDIDK 480
 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNDMLMTVTKKSDIDK 480
 481 QREVTYVWKKFVADRK 496
 472 QREVTYVWKKFVADRK 487

DR EMBL; AF361836; AAK32848.1; -
 DR EMBL; AY124860; AAM70569.1; -
 DR InterPro; IPR002129; Pyridoxal dec.
 DR Pfam; PF0282; pyridoxal dec; I.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 SQ SEQUENCE 493 AA; 56005 MW; 36B3FC93P2978168 CRC64;
 Query Match 81.7%; Score 2135.5; DB 10; Length 493;
 Best Local Similarity 82.4%; Pred. No. 3.3e-153;
 Matches 411; Conservative 32; Mismatches 45; Indels 11; Gaps 2;
 QY 1 MVLHAVSESDVSHSTFASRYVRSISLPRFKWPENSIPKEAAYQIINDELMDGNPRLNI 60
 Db 1 MVLSTVSESDVSHSTFASRYVRSISLPRFKWPENSIPKEAAYQIINDELMDGNPRLNI 60
 QY 61 ASFVTTWPEPCDKLIMSINKNYVDMDEYVPTTELQNRVNMIAHLFNAPLEAEATAVG 120
 Db 61 ASFVTTWPEPCDKLIMSINKNYVDMDEYVPTTELQNRVNMIAHLFNAPLEAEATAVG 120
 QY 121 VGTGSSBAIMLAGLAFKRWQNKRAEGKPVDPKNIIVTGANVQVCWEKFKARYFEVELKE 180
 Db 121 VGTGSSBAIMLAGLAFKRWQNKRAEGKPVDPKNIIVTGANVQVCWEKFKARYFEVELKE 180
 QY 181 VKLSGGYVMDPQQAVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKNKGTGMDTP 240
 Db 181 VNLREDDYVMDPQAVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKNKGTGMDTP 240
 QY 241 IHVDAASGGFIAFLPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRWKNKEDLPEEL 300
 Db 241 IHVDAASGGFIAFLPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRWKTDLPEEL 300
 QY 301 IFHINVLGADOPTFTLNFSGSSQVIAQYQYLIQLRGLGHEGFRVNMENCRNMIIVLREGLEK 360
 Db 301 IFHINVLGADOPTFTLNFSGSSQVIAQYQYLIQLRGLGHEGFRVNMENCRNMIIVLREGLEK 360
 QY 361 TERFNIVSKDEGVPLVAFSLKDSCHTEFEISDMRLRYGWIVPAYTWPNAQHITVLRVY 420
 Db 361 TGRFNIVSKDEGVPLVAFSLKDSCHTEFEISDMRLRYGWIVPAYTWPNAQHITVLRVY 420
 QY 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNDMLMTVTKKSDIDK 480
 Db 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNDMLMTVTKKSDIDK 480
 QY 481 QREVTYVWKKFVADRK 496
 Db 471 QREVTYVWKKL-ETKTN 488

RESULT 9
 Q944L6 PRELIMINARY; PRT; 494 AA.
 ID Q944L6
 AC Q944L6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Atig65960/F12P19_12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Arabidopsis cdNA clones";
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Arabidopsis cdNA clones";
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Arabidopsis ORF clones";
 RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 X!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 X!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC)
 EMBL; AC006532; AAD20099.1; -

RESULT 9
 Q944L6 PRELIMINARY; PRT; 494 AA.
 ID Q944L6
 AC Q944L6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Atig65960/F12P19_12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J., Carninci P., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Arabidopsis cdNA clones";
 RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

C -!- COPACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
C EMBL; AF428294; AAL16126.1; -.
C InterPro: IPR002129; Pyridoxal_dec.
C Pfam; PF00282; pyridoxal_dec.1.
C Decarboxylase; Lyase; Pyridoxal phosphate.
C Q SEQUENCE 494 AA; 56168 MW; 9308E4A2456423C4 CRC64;

Query Match 81.3%; Score 2126; DB 10; Length 494;
Best Local Similarity 80.0%; Pred. No. 1.7e-152;
Matches 403; Conservative 46; Mismatches 43; Indels 12; Gaps 3;

Y 1 MVLHAVSSDVSVHSTFASRYVTSLSRPFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
b 1 MVLTKAT-NDESVCVMFGRYVTRTLPKYEIGENSIPKDAAYQIINDELMLDGNPRLNL 59
Y 61 ASFVTTWMEPECDKLIINSSINKNYVDMDEYVTTTELQNCVMIHLNAPLEBAETAVG 120
b 60 ASFVTTWMEPECDKLIINSSINKNYVDMDEYVTTTELQNCVMIHLNAPLEBAETAVG 119

121 VGTGSSSEAIMLAGLAFKRWKQKRAEGKPVDPKPNIVTGANVQVCWEKFAFYFEVELKE 180
b 120 VGTGSSSEAIMLAGLAFKRWKQKRAEGKPVDPKPNIVTGANVQVCWEKFAFYFEVELKE 179

181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEPEFVKLLDLIVKKNKFTGWDTTP 240
b 180 VNLSEGYVMDPQAAEMVDENTICVAAILGSTLNGEPEFVKLLDLIVKKNKFTGWDTTP 239

241 IHVDAASGGFIAPFLYPELWDFRPLVKSINVSGHKYLAVYAGIGWVIRWKNKEDLPEEL 300
b 240 IHVDAASGGFIAPFLYPELWDFRPLVKSINVSGHKYLAVYAGIGWVIRWRAEDLPEEL 299

301 IFHINYLGDQPTFTLNFSGSSQIAQYQLIRLGHGGRVNVNCRENMIVLREGLEK 360
b 300 IFHINYLGDQPTFTLNFSGSSQIAQYQLIRLGHGGRVNVNCRENMIVLREGLEK 359

361 TERFNIVSKDEGVPLVAFSLKSSCHTEFEISDMRRYGVIPAYTMPNAQHITVLRVY 420
b 360 TERFNIVSKDQGVVAVFSLKSHSFEISEMLRRFGWIVPAYTMPADVQHIITVLRVY 419

421 IRDFSRTLAERLAVIEKVMRDELPSRVHKKISLUGQKSESNDLMTVTKSDIDK 480
b 420 IRDFSRTLAERLAVDISKVIHELDTLPSKSKWNGI-----EGIAENVYKEMKMEK 470

481 Q--RDIITGWKKFVADRKKTSGI 502
b 471 EILMEVIVGWRKFKVGRKQWNGVC 494

RESULT 10
92FPS4
Q9ZPS4 PRELIMINARY; PRT; 500 AA.
C Q9ZPS4
C 01-MAY-1999 (TrEMBLrel. 10, Created)
T 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
T 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
T Putative glutamate decarboxylase.
N AT2G02000 OR AT2G02000/F14H20.7.
N Arabidopsis thaliana (Mouse-ear cress).
C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C Spermatoxyla; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
C eurosids II; Brassicales; Brassicaceae; Arabidopsids.
C NCBI_TaxID=3702;
C SEQUENCE FROM N.A.
C STRAIN=cv. Columbia;
C X MEDLINE=20083487; PubMed=10617197;
C Lin X.; Kaul S.; Rounsley S.D.; Shea T.P.; Benito M.-I.; Town C.D.;
C Fujii C.Y.; Mason T.M.; Bowman C.L.; Barnstead M.E.; Feldblum T.V.;
C Buell C.R.; Ketchum K.A.; Lee J.J.; Rensing C.M.; Koo H.; Moffat K.S.;
C Cronin L.A.; Shen M.; VanAken S.E.; Umayam L.; Tallon L.J.; Gill J.E.;
C Adams M.D.; Carrera A.J.; Creasy T.H.; Goodman H.M.; Somerville C.R.;

RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana."
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- COPACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY)
CC -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
TYRDC).
DR EMBL; AC006532; AAD20093.1; -.
DR EMBL; AK118125; BAC42751.1; -.
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec.1.
KW Decarboxylase; Lyase; Pyridoxal phosphate.
SQ SEQUENCE 500 AA; 56506 MW; 1A47C0D14C6CD3B8 CRC64;

Query Match 79.5%; Score 2080; DB 10; Length 500;
Best Local Similarity 78.8%; Pred. No. 5.4e-149;
Matches 395; Conservative 44; Mismatches 52; Indels 10; Gaps 1;

QY 1 MVLHAVSSDVSVHSTFASRYVTSLSRPFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
Db 1 MVLSTASKSDDSIHSTFASRYVTRNSISRFIPKNSIPKEAAYQIINDELMLDGNPRLNL 60
QY 61 ASFVTTWMEPECDKLIINSSINKNYVDMDEYVTTTELQNCVMIHLNAPLEBAETAVG 120
Db 61 ASFVTTWMEPECDKLIINSSINKNYVDMDEYVTTTELQNCVMIHLNAPLEBAETAVG 120

121 VGTGSSSEAIMLAGLAFKRWKQKRAEGKPVDPKPNIVTGANVQVCWEKFAFYFEVELKE 180
Db 121 VGTGSSSEAIMLAGLAFKRWKQKRAEGKPVDPKPNIVTGANVQVCWEKFAFYFEVELKE 180

181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEPEFVKLLDLIVKKNKFTGWDTTP 240
Db 181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEPEFVKLLDLIVKKNKFTGWDTTP 240

241 IHVDAASGGFIAPFLYPELWDFRPLVKSINVSGHKYLAVYAGIGWVIRWKNKEDLPEEL 300
Db 241 IHVDAASGGFIAPFLYPELWDFRPLVKSINVSGHKYLAVYAGIGWVIRWTKTDLPEEL 300

301 IFHINYLGDQPTFTLNFSGSSQIAQYQLIRLGHGGRVNVNCRENMIVLREGLEK 360
Db 301 IFHINYLGDQPTFTLNFSGSSQIAQYQLIRLGHGGRVNVNCRENMIVLREGLEK 360

361 TERFNIVSKDEGVPLVAFSLKSSCHTEFEISDMRRYGVIPAYTMPNAQHITVLRVY 420
Db 361 TGRFNIVSKENGVPLVAFSLKSSRHRNEFEVAEMLRFRFGWIVPAYTMPADVQHIITVLRVY 420

421 IRDFSRTLAERLAVIEKVMRDELPSRVHKKISLUGQKSESNDLMTVTKSDIDK 480
Db 421 IRDFSRTLAERLAVDFEKVHELDTLPSRVHKKISLUGQKSESNDLMTVTKSDIDK 480

481 Q--RDIITGWKKFVADRKKTSGI 501
Db 471 QREVTAYWKKFVDTKDKNGV 491

RESULT 11
Q9AQU4
ID Q9AQU4 PRELIMINARY; PRT; 501 AA.
AC Q9AQU4;

T 01-JUN-2001 (TrEMBLrel. 17, Created)
 T 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 T 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 E Glutamate decarboxylase (EC 4.1.1.15).
 E GAD.
 S Oryza sativa (Rice).
 C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 C Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 C Ehrhartoideae; Oryzaceae; Oryza.
 X NCBI_TaxID=4530;
 X [1]
 N N
 N N
 P SEQUENCE FROM N.A.
 C STRAIN=cv. Nipponbare; TISSUE=Shoot;
 A Akama K., Akihiro T., Kitagawa M., Takaiwa F.;
 T "Molecular characterization of two genes encoding glutamate
 T decarboxylase from rice (*Oryza sativa*).";
 L Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 C -|- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 C -|- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 C TYRDC).
 R EMBL; AB056062; BAB32870.1; -.
 R EMBL; AB056060; BAB32868.1; -.
 R Gramene; O9A0U4; -.
 R InterPro; IPR002129; Pyridoxal dec.
 R Pfam; PF00282; pyridoxal_dec.i.
 W Decarboxylase; Lyase; Pyridoxal phosphate.
 Q SEQUENCE 501 AA; 56657 MW; E84962C147FFA7A8 CRC64;

Query Match 78.6%; Score 2056.5; DB 10; Length 501;
 Best Local Similarity 78.1%; Pred. No. 3.2e-147;
 Matches 395; Conservative 44; Mismatches 58; Indels 9; Gaps 4;

Y 1 MVLSHAVESD-----VSVSTFASRYVRSLSRFRKWPENSIPKEAAYQIINDELMDGNP 56
 b 1 MVVSAATSDTAQVQYSTFFASRYVRSLSRFRKWPENSIPKEAAYQIINDELMDGNP 60
 Y 57 RNLASFTVWMEPECDKLIIMSSINKNIVDMDEYVVTTELQNRKCNVMIHLNAPLEAE 116
 b 61 RNLASFTVWMEPECDKLIIMDSVKNKYVDMDEYVVTTELQNRKCNVMIHLNAPLEAE 120
 Y 117 TAVGVGTVSSSEAIMLAGLAFKRWKQKAEKGPVDPKNIVTGANVQVCWEKFAFYFV 176
 b 121 TAVGVGTVSSSEAIMLAGLAFKRWKQKAEKGPVDPKNIVTGANVQVCWEKFAFYFV 180
 Y 177 ELKEVKLSEGYVMDPQQAQVDMVDENITCVADILGSLTNGEFEDVKLLNDLLVERKNETG 236
 b 181 ELKEVKLSEGYVMDPVKAVEMVDENITCVAAILGSLTNGEFEDVKLLNDLLVERKNETG 240
 Y 237 WDPPIHVDAAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVIRNKEDL 296
 b 241 WDPPIHVDAAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVIRNKEDL 300
 Y 297 PEELIFHINYLGDQPTFTLNFSGSSQVIAQYQILRLGHEGYNVMNCRNEMVLR 356
 b 301 PEELIFHINYLGDQPTFTLNFSGSSQVIAQYQILRLGHEGYNVMNCRNEMVLR 360
 Y 357 GLEKTRFNIVSKDEGVPVAVFSLKSSCHTEPEISDMRRYGVIVPAYMPPNAQHITV 416
 b 361 GIEATRFEILSKEAGVPLVAVFSLKSSGRYTVFDIIEHLRFRFGVIVPAYMPPNAESHVAV 420
 Y 417 LRVVREDFRSLAERLIVDIEKVMRELDLPSRVHVKISLQEKSESNSDNLMTVTKS 476
 bb 421 LRVVREDFRSLAERLIVSDIVKILHEDLAHSAQVL-KISSAKQQSGDDG--VVTKS 477
 Y 477 DIDKQDIIIGWKKFVADRKKTSIGC 502
 bb 478 VLETEREIFAYRDQV--KKQKIGC 501

RESULT 12
 D 09LSH2
 D 09LSH2 PRELIMINARY; PRT; 494 AA.
 C 09LSH2

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Glutamate decarboxylase.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 CC -|- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 CC TYRDC).
 DR EMBL; AB026646; BAB02870.1; -.
 DR InterPro; IPR002129; Pyridoxal dec.
 DR Pfam; PF00282; pyridoxal_dec.i.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 SQ SEQUENCE 494 AA; 55770 MW; 7985F175E54DF262 CRC64;

Query Match 75.3%; Score 1969; DB 10; Length 494;
 Best Local Similarity 75.2%; Pred. No. 1.3e-140;
 Matches 373; Conservative 56; Mismatches 61; Indels 6; Gaps 3;

QY 1 MVLSHAVESDVSVHSTFASRYVRSLSRFRKWPENSIPKEAAYQIINDELMDGNPRLN 60
 b 1 MVLA-TNSDSDEHLHSTFASRYVRSVAVVFRFKPDPKDAAYQVINDLMDGNPRLN 59
 QY 61 ASFTVWMEPECDKLIIMSSINKNIVDMDEYVVTTELQNRKCNVMIHLNAPLEAEAVG 120
 b 60 ASFTVWMEPECDKLIIMDSVKNKYVDMDEYVVTTELQNRKCNVMIHLNAPLEAEAVG 119
 QY 121 VGTGSSSEAIMLAGLAFKRWKQKAEKGPVDPKNIVTGANVQVCWEKFAFYFEVLKE 180
 Db 120 CQTGSSSEAIMLAGLAFKRWKQKAEKGPVDPKNIVTGANVQVCWEKFAFYFEVLKE 179
 QY 181 VKLSGYYVMDPQQAQVDMVDENITCVADILGSLTNGEFEDVKLLNDLLVERKNETG 240
 Db 180 VKLSGYYVMDPQQAQVDMVDENITCVAAILGSLTNGEFEDVKLLNDLLVERKNETG 239
 QY 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVIRNKEDLPEEL 300
 Db 240 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVIRNKEDLPEEL 299
 QY 301 IPHINYLGDQPTFTLNFSGSSQVIAQYQILRLGHEGYNVMNCRNEMVLRREGLEK 360
 Db 300 VPHINYLGDQPTFTLNFSGSSQVIAQYQILRLGHEGYNVMNCRNEMVLRREGLEK 359
 QY 361 TERFNIVSKDEGVPVAVFSLKSSCHTEPEISDMRRYGVIVPAYMPPNAQHITVLRV 420
 Db 360 TGFNIVSKDEGVPVAVFSLKSSCHTEPEISDMRRYGVIVPAYMPPNAQHITVLRV 419
 QY 421 IREDFRSLAERLIVDIEKVMRELDLPSRVHVKISLQEKSESNSDNLMTVTKSIDDK 480
 Db 420 IREDFRSLAERLIVDIEKVMRELDLPSRVHVKISLQEKSESNSDNLMTVTKSIDDK 475
 QY 481 QDIIITGKKFVADRK 496
 Db 476 E-DITKYWKRLVEHKK 490

RESULT 13
 3LFR4 PRELIMINARY; PRT; 494 AA.
 O8LFR4
 O8LFR4; 01-OCT-2002 (Tremblrel_22, Created)
 I I 01-OCT-2002 (Tremblrel_22, Last sequence update)
 I I 01-MAR-2003 (Tremblrel_23, Last annotation update)
 E Glutamate decarboxylase, putative.
 S Arabidopsis thaliana (Mouse-ear cress).
 C Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 C eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 X NCBI_TaxID=3702;
 N [1]
 N SEQUENCE FROM N.A.
 P Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
 A Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 I "Full-length messenger RNA sequences greatly improve genome
 L annotation."; Genome Biol. 0:0-0(2002).
 L [2]
 N SEQUENCE FROM N.A.
 P Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 A Feldmann K.;
 I "Full-length cDNA from Arabidopsis thaliana";
 L Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 C -! COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 C -! SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 C TVRDC).
 C EMBL: AY084689; AAM61251.1; -
 R InterPro: IPR002129; Pyridoxal dec.
 R Pfam: PF00282; Pyridoxal dec; I.
 R Decarboxylase; Lyase; Pyridoxal phosphate.
 Q SEQUENCE 494 AA; 55804 MW; 7B875CB82EC23E20 CRC64;
 Query Match 75.1%; Score 1965; DB 10; Length 494;
 Best Local Similarity 75.0%; Pred No. 2.7e-140;
 Matches 372; Conservative 56; Mismatches 62; Indels 6; Gaps 3;
 y 1 MVLHAVSESDVSHSTFASRYVRTSLRPFKMPENSIKPEAAQIINDMLDGNRLNL 60
 b 1 MVLAVTNSDSEHLHSTFASRYVRAVVPFQMPDHCMPKDAAYQVINDELMDGNRLNL 59
 y 61 ASFVTTWMEPCDKLIMSSINKNYVDMDEYFVTTTELQNRQVNMIAHLFNALEAEATVG 120
 b 60 ASFVTTWMEPCDKLIMDSYKNKYVDMDEYFVTTTELQNRQVNMIAHFFAPVGEDEAAIG 119
 y 121 VGTVGSSEALMLAGLAFKRWQNKRAEGRPDKPNIVTGANVQVQWCKEFARYFEVELKE 180
 b 120 CGTVGSSEALMLAGLAFKRWQNKRAEGRPDKPNIVTGANVQVQWCKEFARYFEVELKE 179
 y 181 VKLSEGYVMDPQAVDMDVENTICVADILGSTLNGEFEDVKLLNDLLVEKNETGMDPTP 240
 b 180 VKLSEGYVMDPAKAVEMVDENTICVAAAILGSTLTGTEFEDVKQLNDLLAEKNAETGWEPT 239
 y 241 IHVDAASGGFIAPFLPELEWDFRLPVKSNVSGHKYGLVYAGIAGIWMVWRNKEDLPEEL 300
 b 240 IHVDAASGGFIAPFLPDLEWDFRLPVKSNVSGHKYGLVYAGIAGIWMVWRNKEDLPEEL 299
 y 301 IFHINYLGAQDPTFTLNFSKSSQVIAQYQLIRLGHGEGYRNWMCNENMIVLRREGLEK 360
 bb 300 VFHINYLGAQDPTFTLNFSKSSQVIAQYQLIRLGHGEGYRNWMCNENMIVLRREGLEK 359
 y 361 TERPNIVSKDEGVLVAFSLKDSCHTTEFISDMLRRYGIVPAYTWPNAQIHTVLRVY 420
 bb 360 TGKFNIVSKDIGVPLVAFSLKDSKHTVFEIASELRKFGIIPAYTWPADAAQIHTVLRVY 419
 y 421 IREDFSTLAERLVIDEKVMRELDLPSRVHVKISLGQEKSESNDMLVTVTKSDIDK 480
 bb 420 IREDFSRGLADRLTHIQVLKIEGLPSIAHLAA-----AAVSGDDEBKVKTKAMSL 475
 y 481 QRDITIGWKKFVADRK 496
 : ||| : : :
 : ||| : : :

Db 476 E-DITKYWKRDLVEHKR 490
 RESULT 14
 O8RXH0 PRELIMINARY; PRT; 419 AA.
 AC O8RXH0;
 DT 01-JUN-2002 (Tremblrel_21, Created)
 DT 01-JUN-2002 (Tremblrel_21, Last sequence update)
 DT 01-OCT-2002 (Tremblrel_22, Last annotation update)
 DE Glutamate decarboxylase, putative.
 GN ATIG6960;
 OC Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shimm P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -! COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -! SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 CC TVRDC).
 DR EMBL: AY081259; AAL91148.1; -
 DR InterPro: IPR002129; Pyridoxal dec.
 DR Pfam: PF00282; Pyridoxal dec; I.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 SQ SEQUENCE 419 AA; 47639 MW; 7565ADB31A76957F CRC64;
 Query Match 70.6%; Score 1845.5; DB 10; Length 419;
 Best Local Similarity 81.3%; Pred No. 2.3e-131;
 Matches 348; Conservative 36; Mismatches 33; Indels 11; Gaps 2;
 y 77 MSSINKNYVDMDEYFVTTTELQNRQVNMIAHLFNALEAEATVGTVGSSEALMLAGLA 136
 Db 1 MDSINKNYVDMDEYFVTTTELQNRQVNMIAHLFNALEAEATVGTVGSSEALMLAGLA 60
 y 137 FPKWQNKRAEGRPDKPNIVTGANVQVQWCKEFARYFEVELKEVKSLSGYYVMDPQAV 196
 Db 61 FPKWQNKRAEGRPDKPNIVTGANVQVQWCKEFARYFEVELKEVKSLSGYYVMDPQAA 120
 y 197 DMDVENTICVADILGSTLNGEFEDVKLLNDLLVEKNETGMDPTIHDVDAASGGFIAPFLY 256
 Db 121 EMDVENTICVAAAILGSTLNGEFEDVKRLNDLLVKNBETGMDPTIHDVDAASGGFIAPFIY 180
 y 257 PELEWDFRLPVKSNVSGHKYGLVYAGIAGIWMVWRNKEDLPEELIFHINYLGAQDPTFTL 316
 Db 181 PELEWDFRLPVKSNVSGHKYGLVYAGIAGIWMVWRNAEADLPEELIFHINYLGAQDPTFTL 240
 y 317 NFSKSSQVIAQYQLIRLGHGEGYRNWMCNENMIVLRREGLEKTERENIVSKDEGVLV 376
 Db 241 NFSKSSQVIAQYQLIRLGHGEGYRNWMCNENMIVLRREGLEKTERENIVSKDGGVFW 300
 y 377 AFSLKDSCHTTEFISDMLRRYGIVPAYTWPNAQIHTVLRVWRNEDFSTLAERLVID 436
 Db 301 AFSLKDSCHTTEFISDMLRRYGIVPAYTWPADAAQIHTVLRVWRNEDFSTLAERLVID 360
 y 437 IEKVMRELDLPSRVHVKISLGQEKSESNDMLVTVTKSDIDKQ--RDIITGKKKVVAD 494
 Db 361 ISKVLHELDLPSRISKQWGI-----EGIAKVKKEKKEKEILMEIVIGWRKFVKE 411
 y 495 RKTSGIC 502
 ||| : : :
 ||| : : :
 Db 412 RKOANGVC 419
 RESULT 15

Job time : 73 secs

```

AR41
0 Q9AR41 PRELIMINARY; PRT; 500 AA.
1 Q9AR41:
2 01-JUN-2001 (TrEMBLrel. 17, Created)
3 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
4 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
5 Glutamate decarboxylase (EC 4.1.1.15).
6 GAD.
7 Oryza sativa (Rice).
8 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
9 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
10 Ehrhartoideae; Oryzeae; Oryza.
11 NCBI_TaxID=4530;
12 [1]
13 SEQUENCE FROM N.A.
14 STRAIN=cv. Nipponbare; TISSUE=Shoot;
15 Akama K., Akihiro T., Kitagawa M., Takaiwa F.;
16 "Molecular characterization of two genes encoding glutamate
17 decarboxylase from rice (Oryza sativa).";
18 Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
19 -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
20 -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
21 TYRDC)
22 EMBL; AB056063; BAB32871.1; -
23 EMBL; AB056061; BAB32869.1; -
24 Gramene; Q9AR41; -
25 InterPro; IPR002129; Pyridoxal deC.
26 Pfam; PF00282; pyridoxal deC. 1.
27 Decarboxylase; Lyase; Pyridoxal phosphate.
28 SEQUENCE 500 AA; 55615 MW; 0E19E61CCAD5F508 CRC64;
Query Match 70.5%; Score 1844; DR 10; Length 500;
Best Local Similarity 69.6%; Pred. No. 3.9e-131;
Matches 355; Conservative 56; Mismatches 81; Indels 18; Gaps 5;
Y 1 MVLSH-AVSESDVSHSIFASRYVVRTSLPRFKWPNSEIPKEAAVQIINDELMLDGNPEL 58
b 1 MVLTHVEAVEEGSERAAAVFASRYVQDPFRYELGERSISKDAAQIVHDELLDSSPRL 60
Y 59 NLASFVTTWMEPCDKLIMSSINKNVYDMDVEYVTTTELQNRQVNMIAHLFNAPLEEAE 118
b 61 NLASFVTTWMEPCDRLLLEAINKNYADMDEYVTTTELQNRQVNIILARLFNAPVGDGEKA 120
Y 119 VGVGTGSSSAIMLAGLAFKRWKQNKRAEGKVDKPNIVTGANVQVCKEKFARYPEVEL 178
b 121 VGVGTGSSSAIMLAGLAFKRRYQNRKRAAGKPKPNIVTGANVQVCKEKFARYPEVEL 180
Y 179 KEVKLSRGYVMDPQOAVDMVDENTICVADILGSTLNGRPFEDVKLLNDLLVEKNKETGWD 238
b 181 KEVKLTGECYVMDPVKAVDMVDENTICVAALLGSTLTGFEDVRRLNDLLAAKNKKTGWD 240
Y 239 TPIHVDAAAGGFIAPELYPELEWDFRLPVKSNVSGHKYGLVYAGIGWYIMRNKEDLPE 298
b 241 TPIHVDAAAGGFIAPELYPELEWDFRLPVKSNVSGHKYGLVYAGYVWVWVWVWVWVWV 300
Y 299 ELIPIHNYLGNADQPTFTLNFSKSSQVIAQYQQLILGHGEGYRNWNCENMIVLREG 358
b 301 ELIPIHNYLGNADQPTFTLNFSKSSQVIAQYQYQYQYQYQYQYQYQYQYQYQYQYQY 360
Y 359 EKTERFNVSKDEGVPLVAFSLKSDSCHTFEISDMLRYYGWIWVPAWTPPNAACHITVLR 418
b 361 EKUGRFIIISKEEGVPLVAFVFKDGAQAQAFRLSSGLRKYGWIWVPAWTPPAALEHNTVVR 420
Y 419 VVIREDFRTLAERLVIDIEKVMKELD-----ELPSRVHHSILGQEKSESDNLMVTV 473
b 421 VVREDFGRPLAERFLSHVRMALDEMDLAARAVP-RVQLTIBLGPARTAGEBASIRVVK 479
Y 474 KKS-DIDKQRDIITGWKKFYADRKKTSGIC 502
b 480 SEAVFVRKSVPLVAG-----KTKGVC 500

```

GenCore version 5.1.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.
 M protein - nucleic search, using frame_plus_p2n model
 run on: October 22, 2003, 16:47:21 ; Search time 359 Seconds
 (without alignments)
 3750.039 Million cell updates/sec
 title: US-10-006-852-2
 perfect score: 2615
 sequence: 1 MWLSHVSSEDSVSRHSTFAS.....DIITGKKFVDRKKTSGIC 502

scoring table:
 BLOSUM62
 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0
 searched: 1792395 seqs, 1340900451 residues
 total number of hits satisfying chosen parameters: 3584790
 minimum DB seq length: 0
 maximum DB seq length: 2000000000
 post-processing: Maximum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

command line parameters:
 Q=/cgn2_1/USPTO.spool_p/US10006852/runat_21102003_164105_2567/app_query.fasta_1.647
 DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpp -MINMATCH=0.1
 LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
 TRANS=human40.cdi -LIST=45 -DCCALIGN=200 -THR_SCORE=pct -THR_WAX=100
 THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0
 MAXLEN=200000000 -USER=US10006852 @CGN_1_107.@runat_21102003_164105_2567
 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
 FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/prodata/2/pubpna/US05_NEW_PUB.seq:*
 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
 13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
 14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
 15: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
 16: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
 17: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
 \$
 SUMMARIES
 result Query Match Length DB ID Description

 1 2615 100.0 1509 14 US-10-006-852-1 Sequence 1, Appli

2	2607	99.7	1509	10	US-09-938-842A-1810
3	2250	86.0	1785	14	US-10-006-852-15
4	2240	85.7	1705	14	US-10-006-852-11
5	2233	85.4	1745	14	US-10-005-602-1
6	2200	84.1	1479	10	US-09-887-576-820
7	2199	84.1	1771	14	US-10-006-852-13
8	2135.5	81.7	1482	10	US-09-938-842A-937
9	2130	81.5	1665	14	US-10-006-852-3
10	2073	79.3	1497	12	US-10-167-547C-1
11	2014	77.0	1783	14	US-10-006-852-17
12	2007	76.7	1455	10	US-09-887-576-782
13	1936.5	74.1	2121	14	US-10-006-852-7
14	1893.5	72.3	1287	10	US-09-887-576-806
15	1864	71.3	1530	12	US-10-167-547C-3
16	1854	71.3	1590	12	US-10-167-547C-5
17	1854	70.9	1467	10	US-09-887-576-777
18	1836	70.2	2493	14	US-09-727-758-1
19	1834.5	70.2	2493	14	US-10-006-852-5
20	1782	68.1	1946	14	US-10-006-852-9
21	1630	62.3	5160	10	US-09-887-576-443
22	1314.5	50.3	1302	10	US-09-887-576-783
23	1184	45.3	894	12	US-10-167-547C-30
24	1082	41.4	1410	14	US-10-156-761-4639
25	1082	41.4	9025608	14	US-10-156-761-1
26	1051	40.2	1362	14	US-10-156-761-3588
27	1051	40.2	9025608	14	US-10-156-761-1
28	934	35.7	742	10	US-09-966-881-20
29	851	32.5	807	10	US-09-966-881-21
30	620.5	23.7	440	9	US-09-770-444-922
31	545	20.8	387	10	US-09-878-574-1480
32	452	17.3	493	11	US-09-770-961-610
33	438	16.7	280	10	US-09-878-574-7302
34	436.5	16.7	781	11	US-09-989-442-73
35	347	13.3	277	10	US-09-878-574-13198
36	325	12.4	1770	12	US-10-197-073-7
37	325	12.4	1770	12	US-10-053-510-1
38	325	12.4	1770	14	US-10-286-175-7
39	315.5	12.1	477	9	US-09-770-444-28
40	299.5	11.5	1638	12	US-10-053-510-15
41	265.5	10.2	1629	12	US-10-197-073-5
42	265.5	10.2	1629	14	US-10-053-510-3
43	265.5	10.2	1629	14	US-10-286-175-5
44	257	9.8	359	12	US-10-084-843-192
45	257	9.8	359	12	US-10-193-002-187

ALIGNMENTS

RESULT 1
 US-10-006-852-1
 ; Sequence 1, Application US/10006852
 ; Publication No. US20030046732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kinnersely, Alan M.
 ; APPLICANT: Turano, Frank J
 ; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 ; FILE REFERENCE: 7224-65
 ; CURRENT APPLICATION NUMBER: US/10/006.852
 ; PRIOR FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: US 60/246,367
 ; PRIOR FILING DATE: 2000-11-07
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1509
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1509)
 ; OTHER INFORMATION:
 US-10-006-852-1

Alignment Scores: 8.08e-295 Length: 1509
 Red. No.: 2615.00 Matches: 502
 Score: 100.00% Similarity: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Gaps: 14

US-10-006-852-2 (1-502) x US-10-006-852-1 (1-1509)
 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 1 AFGGTGCTCTCCACCGCCGATATCGGAGTGGGACGCTCCGTCACCTCCACATTCGCATCA 60
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 61 CGTTACGTCCTGACTTCACTTCTAGTTCAGATGCGGAAACTCGATTCTTAAGGAA 120
 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 121 GCGGGGTATCAGATCAACCAACGACGAGCTGATGCTTGGCGGAAATCCACGGTTGAACATTA 180
 61 AlaSerPheValThrThrIleMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 181 GCCTCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAATCATCATGTCCTCCATC 240
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 241 AACAGAACTATGTTGACATGGACGAGTACCCCGTCCACCCGAACTTCAGAACCGATGT 300
 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
 301 GTGAACATGATGACATCTATTTCATGATGACCGTGTAGAAGGCGGAGACCCCGCTCGGA 360
 121 ValGlyThrValGlySerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 361 GTAGGAACCGTTGGATCATCGGAGGCCATAAATGTTGGCCGCTTGGCCCTTCAAGCGTAAA 420
 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 421 TGGCAGAACAAAGCCAAAGTGAAGCAAAACCGTTCGATAAAACCCCAAACTTGCACCCGGA 480
 161 AlaAsnValGlnValCysTrpGluLysPheAlaAaArgTyrPheGluValGluLeuLysGlu 180
 481 GCGCAATGTTCAAGTGTGGGAGAAATTCGTTAGTACTTTGAGGTTGAACTTAAGGAA 540
 181 ValLysLeuSerGluGlyTyrValMetAspProGlnAlaValAspMetValAsp 200
 541 GTGAATTTGAGTGAAGATACATATGATGATGACCCCTCAACAAGCTTGTGATGGTTGAT 600
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 601 GAGAACACCATTTGTTGGGACATCTTGTGTTCCACTTTAATGGAGAAATTCGAAGAT 660
 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 661 GTTAAACTCTTGAACGATCTTTGGTTCGAAAGAACAAAGAACCCGATGGGATACACCA 720
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 721 ATCCAGTGGATGGCAAGTGGAGATTCATTCACCGCTTTTGTGATCCGGAATTTGGAA 780
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 781 TGGGACTTTAGACTTCCCTTGGTCAAGAGTATCAATGTGAGTGGTCCACAAATGGACTT 840
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuLeu 300
 841 GTGTACCGAGGATTTGGTGGTGGTATCTGGAGAACAAAGAGGATTTGCTTGGAGACTC 900
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 901 ATCTTCCATCAATATCTTGGTGGTGGACCAACCCACTTTTACTCTCAATTTCTCCAAA 960

321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 961 GGTTCAAAGTCAAGTTCATTCCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 1020
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 1021 AGAAATGTTGATGGAGAAATTCAGAGAGAAATATGATCTCTTAAAGGAGGACTTGAAG 1080
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 1081 ACAGAAAGTTCACATCGTCTCAAGAGGACGAGGAGTGCACCTTTCGCTTTCCTCTTG 1140
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTyr 400
 1141 AAAGATACAGCTCTCACATGAGTTCGAATCTCCGACATCTCCGACATCTCCGACATCT 1200
 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1201 ATAGTCCGCGCTACACAAATGCCCTCCAAATGCACAAACATCACTGTTCTTCGTGTG 1260
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1261 ATCAGAAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGT 1320
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1321 ATGCGTGAAGTTCGATGAGCTTCTTCGAGAGTGTATCAGAAATATCACTTGGACAAG 1380
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1381 AAGAGTGAATCTACAGCGATAACTTGTATGTCACCGTGAAGAGAGCGGATATCGACA 1440
 481 GlnArgAspIleIleThrGlyTrpLysPheValAlaAspArgLysLysThrSerGly 500
 1441 CAGAGAGATATCATCACTGGCTGGAGAAAGTTTTCGCCGACAGGAAAGAGACGAGTGG 1500
 501 IleCys 502
 1501 ATCTGC 1506

RESULT 2
 US-09-938-842A-1810
 ; Sequence 1810, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Krebs, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1810
 ; LENGTH: 1509
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1810

Alignment Scores: 6.94e-294 Length: 1509
 Red. No.: 2607.00 Matches: 501
 Score: 100.00% Similarity: 99.80% Conservative: 0
 Percent Similarity: 99.80% Mismatches: 1
 Best Local Similarity: 99.80%

```

ery Match: 99.69% Indels: 0 Gaps: 0
: 10
-10-006-852-2 (1-502) x US-09-938-842A-1810 (1-1509)
1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
1 ATGGTCTCCACGGCCGATCGGAGTCGGACGTCCTCCGTCACATCCACATTCGCATCA 60
21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
61 CGTTACGTCGGTACCTTCTAGTTCAGATGCGCGAATCCGATTCCTTANGRA 120
41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
121 GCGGCGTATCAGATCATCAACGACGCTGATGCTTGACGGAAATCCACGGTTGAACIT 180
61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
181 GCCTCCTTTGTACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
241 AACAGAACTATGTTGACATGAGCGAGTACCCTCCGTCACCCGAACTTCAGAACCGATGT 300
101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
301 GTGACATGATGCATCTAATCAATGACCGTTAGAAAGCGGAGACCCGCTCGGA 360
121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
361 GTAGAAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGTTTGGCCCTCAAGCGTAAA 420
141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
421 TGGCAGAACAGCGCAAGCTGAGCCAAACCCCTCGATTAACCCCAACTTGTCCCGGA 480
161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
481 GCCAATGTTCAAGTGTGGGAGAAATTCGTTAGGTACTTTGAGGTTGAACTTAAGGAA 540
181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
541 GTGAATGATGAGTAAAGGACTATGATGATGGACCCCTCAACAAGCTGTTGATATGGTTAT 600
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
601 GAGAACCACTTGTGTGGGCCATCTTGGTTCCACTCTTAATGGAGAAATTCGAAGAT 660
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
661 GTTAAACTCTTGAACGATCTTGTGGTGGAAAGAAACAAAAGAAACCGGATGGATACCA 720
241 IleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
721 ATCCAGTGGATGGGCAAGTGGAGATTCATTCACCGTTTTTGTATCCGGAATTTGHA 780
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
781 TGGGACTTTAGACTTCCCTTGGTGAAGAGATCAATGTGAGTGGTCCAAAGATGAGACTT 840
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuLeu 300
841 GTGTACCGAGGATGGTGGTTCATCGGAGAAACAAAAGAGATTTGCTCGAGGAATTC 900
301 IlePheHisLysLeuTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
901 ATCTTCCATATCAATTAICTTGGTGTGCAACCCACCCACTTACTCTCAATTTCTCCAA 960
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
961 GGTTCAAGTCAAGTCAITGCTCAATACCACTATCCGATTCGATTTGGCCACGAGGGTTAC 1020

```

```

341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
1021 AGAAATGTGATGGAGAATTCAGAGAGAAATATGATCGTCTTAAGGGAGGACTTGAGAAG 1080
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
1081 ACAGAAAGGTTCAACATCGTCTCAAGGACGAGGAGTGCCTTGTCTTCTCTCTTG 1140
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTrp 400
1141 AAGATAGCAGCTGTCCACTGATTCGAAATCTCCGACATGCTTCGAGGATGATGATG 1200
401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
1201 ATAGTCCGCGCTACCAATGCTCCAAATGCACAAACATCACTGTTCTTCGTGTGGTT 1260
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
1261 ATCAGAAAGATTTTCGAGAACACTCGCTGAGACTTGTGATCGATATAGAGAAAGTG 1320
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1321 ATCGGTGAGCTCGATGAGCTTCTTCGAGAGTGTATTCACAAAATATCACTTGGACAAG 1380
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1381 AAGAGTGAATCTAACACCGATACTTGTATCGTCCGCTGAGAGGAGCGGATATCGCAAG 1440
481 GlnArgAspIleIleThrGlyLysPheValAlaAspArgLysLysThrSerGly 500
1441 CAGACAGATATCACTCTGGCTGGAGAAATTTGTCCCCGACAGGAAAGACGAGTGGT 1500
501 IleCys 502
1501 ATCTGC 1506
RESULT 3
US-10-006-852-15
; Sequence 15, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
; APPLICANT: Kinnersely, Alan M.
; APPLICANT: Turano, Frank J.
; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
; FILE REFERENCE: 7224-65
; CURRENT APPLICATION NUMBER: US/10/006,852
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/246,367
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 15
; LENGTH: 1785
; TYPE: DNA
; ORGANISM: Petunia x hybrida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(1571)
; OTHER INFORMATION:
US-10-006-852-15
Alignment Scores: 4.19e-252 Length: 1785
Pred. No.: 2250.00 Matches: 427
Score: 92.64% Conservativeness: 39
Best Local Similarity: 84.89% Mismatches: 33
Query Match: 86.04% Indels: 4
DB: 14 Gaps: 2
US-10-006-852-2 (1-502) x US-10-006-852-15 (1-1785)
QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20

```


72 ATGGTTCATCAAGACAGTGTCCGAGCGCATGTCCCAITTCACCTCCACGTTTGGCTTCT 131
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 132 CGATATGTTCCGAACTTCTCTCCAGGTTTAAATGCCAGATAAATTCGATACCAAGAA 191
 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 192 GCAGCATATCAGATATAAATGATGAATGATGATGATGATGATGATGATGATGATGATG 251
 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 252 GCTTCTTTTGTACAACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 311
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 312 ARCAAGAACTATGTTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 371
 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
 372 GTAAACATGATAGCTCATTTGTTTAAATGACCACTTGAAGATGGAGAAACTGCAGTTGGA 431
 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 432 GTTGGAACTGTGGATCCCTCTGAGCCATATGCTTGTCTGGATTAGCTTTCAAGAGAAA 491
 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 492 TGGCAGACAAATAAGAAAGCCCAAGGCAAAACCTGTGCAAGCCCAACATTTGTTACTGGT 551
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 552 GCAAAATGTCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAGAA 611
 181 ValLysLeuSerGluGlyTyrValMetAspProGlnIleAlaValAspMetValAsp 200
 612 GTPAAGCTTAGTGAAGAACTATGATGATGATGATGATGATGATGATGATGATGATGAT 671
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGlnPheGluAsp 220
 672 GAAACACCAATTTGCTAGTGTCTATCTTAGTTCACCTCAATCGGAAATTTGAAGAC 731
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrClyTrpAspThrPro 240
 732 GTPAAGCCCTTGAATGATCTTCTGGTGGAGAAACAAAGAAACCCGGGAGCACTCCA 791
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 792 ATTCAATGGGATCAGCAAGTGGTGGATTTATTGCAACCGTTCATTTACCAGAGCTTGAG 851
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 852 TGGGACTTTAGATGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuGlu 300
 912 GTCTATGCTGGTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 971
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 972 ATCTTCCACATTAATATCTTGGTGTGATCAACCTTACTTCACTCAACTTTTCTAAA 1031
 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 1032 GGTTCTAGCCAAAGTAAATGCTCAATATATACCAACTTATTCGCTTGGGTTATGAGGGTAC 1091
 341 ArgAsnValIleMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 1092 AAGAATCTGATGGGAATTCGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1151
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 1152 ACAGAAAGATTCACATAAATCTCCAAAGAAATTTGGAGTACCTTTTAGTACCTTCTCTCT 1211

381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTrp 400
 1212 AAAGACAACAGGCAACACACAGGTTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271
 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1272 ATGTGTTCTGCATATACTATGCCCAACGCAACACATPACAGTTCTCAGAGTTGTG 1331
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1332 ATCAGAGAAATTTCTCCGTTACCTTGCAGAACGACTGGTAAAGAGACATCGAAAAGTC 1391
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1392 CTTTCATGAATTCACACACTCCCTGCACGCTGCAATGCTAAGCTCGCTGCGCCGAGGAG 1451
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1452 CAGGCGGCTGCGAATGCGCAGCGAGGTGCAT-----AAGAAAACAGATAGCGAAGTG 1502
 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
 1503 CAGTTGAGATGATACTGCATGGAAAGATTTGTTGAAGAAAAGAGAGAAAGACTAAT 1562
 500 GlyIleCys 502
 1563 CGAGTTGT 1571

RESULT 4
 US-10-006-852-11
 ; Sequence 11, Application US/10006852
 ; Publication No. US20030046732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kinnersely, Alan M.
 ; APPLICANT: Turano, Frank J.
 ; TITLE OF INVENTION: Methods for regulating Plant GABA Production
 ; FILE REFERENCE: 7224-65
 ; CURRENT APPLICATION NUMBER: US/10/006,852
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: US 60/246,367
 ; PRIOR FILING DATE: 2000-11-07
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 11
 ; LENGTH: 1705
 ; TYPE: DNA
 ; ORGANISM: Nicotiana tabacum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (71)..(1558)
 ; OTHER INFORMATION:
 US-10-006-852-11

Alignment Scores:
 Pred. No.: 5,72e-251 Length: 1705
 Score: 2240.00 Matches: 428
 Percent Similarity: 92.05% Conservative: 35
 Best Local Similarity: 85.09% Mismatches: 32
 Query Match: 85.09% Indels: 8
 DB: 14 Gaps: 3

US-10-006-852-2 (1-502) x US-10-006-852-11 (1-1705)

Oy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 71 ATGGTTCGTCGAAAGAGCGTCCGAAAGTGCATCCACTCCACTTCGCTTCC 130
 Oy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 131 CGATATGTTCCGAACTTCTCTCCAGGTTTAAATGCCAGATAAATTCGATACCAAGAA 190
 Oy 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60

```

191 GCAGCATCATCAAATCATAAATGATGAGCTTATGTTAGATGGAATCCAAAGACTAAATTTA 250
/
61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
/
251 GCATCTTTGTGCAACATGGATGGAACCAAGAGTGTAACAACAGTATGATGATGATCCCAAT 310
/
81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuInAsnArgCys 100
/
311 AACCAAGAAATACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 370
/
101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
/
371 GTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 430
/
121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
/
431 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGGATTAGCTTTCAGAGAAA 490
/
141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
/
491 TGGCAAAATAAATAAGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATAATTTCTCACT 550
/
161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
/
551 GCCAATGTCAGGCTGTTGGAGAAATTTGCAGGTATTTTGAAGTGGAGCTTAAAGAA 610
/
181 ValLysLeuSerGluGlyTyrValMetAspProGlnAlaValAspMetValAsp 200
/
611 GTRAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
/
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
/
671 GAGAACCAATTTGTSTAGTCTGCTTCTGGTTCACACTCAATGGTGAATTTGAAGAT 730
/
221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTyrAspThrPro 240
/
731 GTTAAAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
/
241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
/
791 ATTCATGTGGATGAGCAAGTGGTGGATTTATTCACCACTTCTTATCCAGAGTTGAA 850
/
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyIleu 280
/
851 TGGACITTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910
/
281 ValTyrAlaGlyIleGlyTyrValIleTyrArgAsnLysGluAspLeuProGluLeu 300
/
911 GTTATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970
/
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
/
971 ATCTTCCACATTAATATCTTGTGCTGATCAACCTTCTTCTTCTTCTTCTTCTTCTT 1030
/
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
/
1031 GGTTCAGCCAGTAAATGCTCAATATACCACTTATTCGCTTGGTGGTGGTGGTGGTGG 1090
/
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
/
1091 AAGAAATGTTAGGAAATTTCAAGAAAATGCAAGGGTACTAAGAGAGAGACTTGA AAAA 1150
/
361 ThrGluArgPheAsnIleValSerLysAspGluValProIleuValAlaPheSerIleu 380
/
1151 AGTGGAAATTAACATPAATCCAAAGAAATTTGGAGTTCCATGATGATGATGATGATGAT 1210
/
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTyr 400
/
1211 AAAGACAACAGTCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1270
/
401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
/

```

```

1271 ATTATTCCTGCATATACTATGCCACCAAAATGCTCAACATGTCTCACAGTTCTCAGAGTTGTC 1330
Qy
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleHisPileGluLysVal 440
Db
1331 ATTAGAAGAATTTCTCCGTACACTCCGCGGACGACTGGTAAATAGACATTTGAAAANGTC 1390
Qy
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
Db
1391 CTCACCAGCTAGACACACTTCGGCGAGGGTCAACGCTAAGCTAGCCGTG----- 1441
Qy
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysSerAspIleAspLys 480
Db
1442 --GCCAGGCGGAATGGCAGCGCGTGCAAT-----AAGAAAACAGATAGAGAAGTG 1489
Qy
481 GlnArgAspIleIleThrGlyTyrLysLysPheValAlaAsp---ArgLysLysThrSer 499
Db
1490 CAGCTTGAGATTACTACTGCATGGAGAAATTTGTCGTGATPAAGAAAGAAAGACTAC 1549
Qy
500 GlyIleCys 502
Db
1550 GGAGTTTGT 1558
RESULT 5
US-10-005-602-1
; Sequence 1, Application US/10005602
; Publication No. US20030110530A1
; GENERAL INFORMATION:
; APPLICANT: Barry Shelp
; APPLICANT: Alan Bown
; TITLE OF INVENTION: TRANSGENIC PLANTS HAVING REDUCED
; TITLE OF INVENTION: SUSCEPTIBILITY TO INVERTEBRATE PESTS
; FILE REFERENCE: P84US3
; CURRENT APPLICATION NUMBER: US/10/005,602
; CURRENT FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1745
; TYPE: DNA
; ORGANISM: tobacco plant
US-10-005-602-1
Alignment Scores:
Pred. No.: 3.89e-250 Length: 1745
Score: 2233.00 Matches: 427
Percent Similarity: 91.85% Conservative: 35
Best Local Similarity: 84.89% Mismatches: 33
Query Match: 85.39% Indels: 8
DB: 14 Gaps: 3
US-10-006-852-2 (1-502) X US-10-005-602-1 (1-1745)
Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 31 ATGTTCTGTCCAGACAGCGGTGGAAAGTGAGCTTCCATCCACTCCACTTCGTTCC 90
Qy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 91 CGATATGTTCCGAACCTTCTTCCAGGTTTAAAGATGCAAGAAATTCATATCCAAAGGAA 150
Qy 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 151 GCAGCATATCAGATTAATAATGATGAGCTTATGTTAGATGGAAATCCAAAGCTTAAATTTA 210
Qy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
Db 211 GCATCTTTCTGATCAACATGGATGGAGCCAGAAATGTAATACGTTAATGATGATTCAT 270
Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
Db 271 AACAGAACATCGTTGACATGGATGATGATGATGATGATGATGATGATGATGATGATG 330
Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120

```

Db 1402 ---GCCAGGGCAATGCCAGCGGCTGCAT-----AAGRAAACAGATAGAGAACTG 1449
 Qy 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
 Db 1450 CAGCTAGAGATTACTACTGCATGGAAGAAATTTGCTGCTAATAAGAAAGAAAGACTAAT 1509
 Qy 500 GlyIleCys 502
 Db 1510 GGAGTTTGT 1518
 RESULT 6
 US-09-887-576-820
 ; Sequence 820, Application US/09887576
 ; Patent No. US20020144047A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Budworth, P.
 ; APPLICANT: Brown, D.
 ; APPLICANT: Chang, H.
 ; APPLICANT: Zhu, T.
 ; APPLICANT: Han, B.
 ; APPLICANT: Wang, X.
 ; APPLICANT: Cooper, Bret
 ; TITLE OF INVENTION: Promoters for regulation of plant expression
 ; FILE REFERENCE: 1360.001US1
 ; CURRENT APPLICATION NUMBER: US/09/887,576
 ; CURRENT FILING DATE: 2001-06-25
 ; PRIOR APPLICATION NUMBER: US 60/213,848
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 60/214,087
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 60/258,692
 ; PRIOR FILING DATE: 2000-12-29
 ; NUMBER OF SEQ ID NOS: 875
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 820
 ; LENGTH: 1479
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 US-09-887-576-820
 Alignment Scores: 2.13e-246 Length: 1479
 Score: 2200.00 Matches: 420
 Percent Similarity: 90.84% Conservative: 35
 Best Local Similarity: 83.67% Mismatches: 37
 Query Match: 84.13% Indels: 10
 DB: Gaps: 2
 US-10-006-852-2 (1-502) x US-09-887-576-820 (1-1479)
 Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 1 ATGGTGTCTCCAAAGCGGCTCCGAGAGTGACATGTCGCTGCACCTCCACCTTCGCTCC 60
 Qy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 61 CGTACGTCCTCCGCGCTCCCTCCCAAGGTACCGATGCCGAGAACTCGATCCCGAAGAG 120
 Qy 41 AlaIaIyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 Db 121 GCGGCGTACCAGATCATCAACAGCAGCTGATGCTGGACGCAACCCGCGCTGAACCTG 180
 Qy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 181 GCGTGTCTCCACCCGCTGGATGGAGCCCGAGTCCCAAGCTCATATGGCCGCGCATC 240
 Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
 Db 241 AACRAAATCTAGCTGACATGAGCAGATGCCCTCCACCCAGCTCCAGAACCCGCTGC 300
 Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 Db 301 GTGAACATGATCGGCACCTTTCACGCGCGCTCCGGGAGGAGACGGCGGTGGC 360

331 GTAAATATGATAGCTCATTTGTTTAAATGACACCTTGGAGATGGAGACTCACTTGG 390
 121 ValGlyThrValGlySerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 391 GTTGGAACTGTGGATCCCTGMAAGCTATATGCTTGTGCTGGATTAGCCCTTAAAGAAAA 450
 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 451 TGGCAAAATAAATGAAAGCCCAAGCAAGCCCTTTGATAAGCCCAATATCTGCACCCGCT 510
 161 AlaAsnValGlnValCysTrpGluLysPheAlaAspTyrPheGluValGlnLeuLysGlu 180
 511 GCTAAATGTCAGGTGTGTTGGGAGAAATTTGCAAGGTAATTTTGAAGTGGAGTTGAAAGAA 570
 181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 571 GTPAAATTTGAGTGAATGATGATGATGGACCTGAGAAAGCTGTGAAATGTTGGAT 630
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 631 GAGAAATACCAATTTGCTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAAT 690
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 691 GTTACGCTTTGAAATGACCTTTGATTTGAGAGAAACAAAGAACCCGGTGGACACTCA 750
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 751 ATTCAATGGATGCAAGTGGTGGATTTATTTGCACCACTTCTTATCCAGACTTGA 810
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 811 TGGGACTTTAGATGTCATTTGGTGAAGAGTATTAATGTGAGTGGTTCACAAATATGGTCT 870
 281 ValTyrAlaGlyIleGlyTrpValIleTrpAsnLysGluAspLeuProGluLeuGlu 300
 871 GTCATGCTGGTATTTGGTGGCCATTTGGAGAAATGAGAAAGACTTGCCTGATGAACIT 930
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 931 ATTTTCCACATCAATTACCTTGGTGTGATCAACCTTCTCACTTCACTTCTCAAA 990
 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 991 GGTCTAGCCAAAGTAAATGCTCAATATACCAACTTATTTCGCTGGTGGTGGTGGTAC 1050
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValIleArgGluGlyLeuGluLys 360
 1051 AAGAATGTTATGGAGATTTGCAAGAAATGCAAGGGTATTAAGAGAAAGAAATGNAAAA 1110
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 1111 AGTGAAGATTTCAACATAATCTCCAAAGAAATTTGGAGTTCCTTAGTCAATTTCTCT 1170
 381 LysAspSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
 1171 AAGAACACAGTCAACCAATAGTTCGAATTTCTGAACCTTCTGAAGATTTGGATGG 1230
 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleIleThrValLeuArgValVal 420
 1231 ATTGTTCTGCATATACTATGCCCAACCAATGCTCAACATGTTACACTTCTCAGAGTTGTC 1290
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1291 ATTAGAAAGATTTCTCCCGCACACTAGCCGAGCGACTGGTAATAGACATTTGAAAAAGTC 1350
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1351 CTCACAGGTAGACACACTTCGCGGAGGGTCAACGCTAAGTACCCGCTG----- 1401
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480

121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 361 GTGGCAAGGTCGTTGCGTGGAGGCAATCATGCTGGCCGGCTGGCTTCAAGCGGGG 420
 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 421 TGGCAGAACAAAGCGCAAGCGGAGGCGGAGCGGTTTCGCAAGCCCAACATCATCCCGC 480
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 481 GCCAAGCTGCAGGTGCTGGGAGAGTTCGCCCGCTACTTCGAGGTGAGCTCAAGGAG 540
 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 541 GTGAAGCTCCGCGAGGCTACTACGTCATGGACCCCGAGAGCCGTCGACATGTCACAC 600
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 601 GAGAACACCATTCGCTGCGCGCCATCTCCGGTCCACCCTCAACGGCGAGTTCGAGGAC 660
 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 661 GTCACGCTACTCAACGACCTCTCCGACAGAGAACAGAGACTGGGTGGAGAGCGCG 720
 241 IleHisValAspAlaIleSerGlyGlyPheIleAlaProPheLeuTyrProGlnLeuGlu 260
 721 ATCCAGCTGGACCGCGAGCGCGGGTTCATCGCGCGTTCCTGTACCCCGGAGTGGAG 780
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 781 TGGACTTCGGCTGCGTGGTGAAGACATCACTGAGGCGGTCAAGTACCGGGCTC 840
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuLys 300
 841 GTCTACCGCGGCTCGCTGGTGCATCTGGCGCAACAAGAGGACCTGCCCGAGGACTC 900
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 901 ATCTTCCACATCAACTACCTCGCGCGCAGCCAGCCACCTTACCCTCAACTTCTCCAA 960
 321 GlySerSerGlnValIleAlaGlnTyrTyrGluLeuIleArgLeuGlyHisGluGlyTyr 340
 961 GGCTCCAGCCAGTTCATCGCCAGTACTACAGCTCATCCGCCACCGCTTCGAGGGGTAC 1020
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 1021 AGAACATCATGGAGAACTGCCACGAGAACCGCATGTTGCTGAAGGAAGGGCTGGTGA 1080
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 1081 ACCGGAGGTTGACATCGTCCAGGAGAACAGGGGTGCGCTGGTGGCTTCGCTC 1140
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTyrP 400
 1141 AAGGACCGGACCGGCACGAGTTCGAGATCTCCGACATGTCGCGCTTCGCGCTGG 1200
 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1201 ATCGTCCGCGGTACACATTCGCGCGCAGCCGACGCGCAGCATCAGCTGTCGCGTGC 1260
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1261 ATCCGGAGGAGTTACGCCACCCCTCGCCAGCGCTCTGCTCGACATCGAAGAGGTTG 1320
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1321 ATGTPACCAGCTCGACCGGCTCCCTCCAGCTCATSCCCCGCTGCGCGCGCGCGCG 1377
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1378 -----CTGCTGGTGTCCGCAAGGATCGGACTCGAGACG 1413

481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
 1414 CAGCGTGGTGGAGGCGTGGAGAGAGTTCGTGCTC---GCCAAGAGGACCAACGGC 1470
 501 IleCys 502
 1471 GTCIGC 1476
 RESULT 7
 US-10-006-852-13
 ; Sequence 13, Application US/10006852
 ; Publication No. US20030046732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kinnersely, Alan M.
 ; APPLICANT: Turano, Frank J.
 ; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 ; FILE REFERENCE: 7224-65
 ; CURRENT APPLICATION NUMBER: US/10/006,852
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: US 60/246,367
 ; PRIOR FILING DATE: 2000-11-07
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1771
 ; TYPE: DNA
 ; ORGANISM: Nicotiana tabacum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (67)..(1554)
 ; OTHER INFORMATION:
 US-10-006-852-13
 Alignment Scores:
 Pred. No.: 3.7e-246 Length: 1771
 Score: 2195.00 Matches: 423
 Percent Similarity: 90.85% Conservative: 34
 Best Local Similarity: 84.10% Mismatches: 38
 Query Match: 84.09% Indels: 8
 DB: 14 Gaps: 3
 US-10-006-852-2 (1-502) x US-10-006-852-13 (1-1771)
 QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 DB 67 ATGTTCTGTCCAGACAGCGCTCGAAAGTACGCTCCGFTCACTCCACTTTCGCTCC 126
 QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 DB 127 CGAATGTTCCAACTTCTCCAGGTTTAAATGCCAGAGAATCAATACCAAGGAA 186
 QY 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 DB 187 GCAGCATATCAGATATAAATGATGAGCTTATGTATGATGGAAATCCAAAGCTAAATTA 246
 QY 61 AlaSerPheValThrThrProMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 DB 247 GCAUCTTTCGTTACAACTGATGGATGGCCAGAAAGTAACTACCTTAATGATGGATTCCATT 306
 QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 DB 307 AACAAAGAACTACGTTGACATGATGAATACCTCTAAACCCTGAGCTTCAGAAATCGATGT 366
 QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 DB 367 GTAAATATGATAGTCAATTTGTTAATGCACCCTTGGAGATGGAGAGACTGCGAGTTGGA 426
 QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 DB 427 GTTGAACCTGTTGATCCTCTCAAGCTATTAATGCTTGTGATGCTTAAAGCTTAAAGAAA 486
 QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160

487 TGGCAAAATAAATAAGAAAGCCCAAGCAAGCCCTTTGATAAAGCCCAATAATGTCCACCGGT 546
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 547 GCTAATGTCAGGCTGTTGGGAGAAATTTGAGAGTATTTTGAAGTGGAGTTGAAAGAA 606
 181 ValLysLeuSerGluGlyTyrValMetAspProGlnAlaValAspMetValAsp 200
 607 GTAAAATTTGAGTCACTATGATGATGAGCCCTGAGAAAGCTGTGAAATGCTGGAT 666
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 667 GAGAAATACCAATTTGCTGCTACTTCTAGGTCAACACTCAATGGTGAATTTGAAGAT 726
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspTrpPro 240
 727 GTTAAAGCTTTGATGACCTTTTGTATGAGAGAACAAAGAACCCGGTGGACACTCCA 786
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 787 ATTCAATGTTGATGACCAAGTGGTGGATTTATGACCACTTCTTTATCCAGAGCTTCAA 846
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 847 TGGGACTTTAGATTTGCTTGGAGAGAGTATTAATGTAGTGGTCCAAATATGCTCTT 906
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeu 300
 907 GTCATGCTGTTATGTTGGGCCAATTTGAGGAAATAGGAAGACTTGCCTGATGAACCTT 966
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 967 ATTTTCCACATCAATACCTTGGTGGTGTATCAACCTTCTCACTCAACTTCTCTAAA 1026
 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 1027 GGTTCTAGCCCAAGTAAATGCTCAATATACCAACTTATTCCTGGTGGTTTTGAGGGTTAC 1086
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuLys 360
 1087 AAGAATTTATGAGAAATTTCAAGAAATGCAAGGGTATTAAGAGAAAGAAATTAAGAAA 1146
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 1147 AGTGGAAAGATTCACATAATTCCTCAAGAAATTTGAGTTCCTTCTAGTACTTTCTCTT 1206
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTrp 400
 1207 AAAGACACAGTCAACAAATGAGTTGAAATTTCTGAAACTTAAAGAAATTTGAGTGG 1266
 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1267 ATGTTCTGGCATATACTATGCCACCAATGCTCAACATGTCACAGTCTCAGAGTTGTC 1326
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1327 ATTAGAGAAATTTCTCCCGCACATAGCGGAGCGACTGTAATAGACATTTGAAAGATC 1386
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1387 TTCACCGGATGACACACTTCCGGCGAGGGTCAACGCTAAGCTACCCGTG----- 1437
 461 LysSerGlnAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1438 ---GCCAGGCGAATGCCAGCGCGCTGCTP-----AAGAAACAGATAGAGAAGTG 1485
 481 GluArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
 1486 CAGCTAGAGATTACTACTGATGGTTGAAATTTGTTGCTGTATAAGAAAGAAAGACTAAT 1545
 500 GlyIleCys 502
 1546 GGAGTTGT 1554

RESULT 8
 US-09-938-842A-937
 ; Sequence 937, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 937
 ; LENGTH: 1482
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-937

Alignment Scores: 7.22e-239 Length: 1482
 Pred. No.: 2135.50 Matches: 411
 Score: 88.78% Conservative: 32
 Percent Similarity: 82.36% Mismatches: 45
 Best Local Similarity: 81.66% Indels: 11
 Query Match: 10 Gaps: 2
 US-10-006-852-2 (1-502) x US-09-938-842A-937 (1-1482)
 Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 1 ATGGTTTTGTCTAAGACAGTTTTCCGAAATCTGATGCTCAATCCATCACTTTTGGCTCT 60
 Qy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 61 CGTTACGTCGCAACTCTCTTCCACGATTCGAAATGCTGAGAACTCAATCCCAAGAA 120
 Qy 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 Db 121 GCAGCTTACCAAAATCAACAGCAGAGCTAATGCTCCGATGGTAACCCAGGGCTGAACCTA 180
 Qy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 181 GCTTCTTCTGTCACCATGATGGAGCCAGAAATGTCAAGCTCATGATGGAGTCCATC 240
 Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
 Db 241 AACAAAGAACTACGTCGACATGGAGAGTACCCCTGTCACTGAGCTTCCAGAACCGAGT 300
 Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 Db 301 GTTAAATGATGACCGCTCTTTCACCGCCCGCTTGGTGACGGTGAAGCTGCGGTTGGT 360
 Qy 121 ValGlyThrValGlySerSerGluAlaIleValMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 361 GTTGGCACCGTCCGATCGTGGAGCGGATATGTTGGCCGGTTTGGCTTTTAAAGACAA 420
 Qy 141 TrpGlnAsnLysArgCysAlaGluGlySerProValAspLysProAsnIleValThrGly 160
 Db 421 TGGCAGAAATAAGCTAGAGCCCAAGGGCTTCTTATGATGAAGCCCAATATCGTAACCGGT 480
 Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 481 GCTAATGTTCCAGGTTTCTGGAGAAATTCGCAAGGTTATTTCCGAAATGGAGCTTAAGGAA 540

```

; FILE REFERENCE: 7224-65
; CURRENT APPLICATION NUMBER: US/10/006,852
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/246,367
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(1498)
; OTHER INFORMATION:
US-10-006-852-3

Alignment Scores: Length: 1665
Fred. No.: 3.8e-238 Matches: 404
Score: 2130.00 Conservative: 46
Percent Similarity: 89.29% Mismatches: 42
Best Local Similarity: 80.16% Indels: 12
Query Match: 81.45% Gaps: 3
DB: 14

US-10-006-852-2 (1-502) x US-10-006-852-3 (1-1665)
Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 17 ATGTTTTGACAAAAACCGCAACG---AATGATGAATCTCTGCACCATGTTCCGGATCT 73
Qy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 74 CGCTATGTTCCACTACATCTCCCAAGTATGAGATTCGGTGAAGATTCGATACCGAAAGAC 133
Qy 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 134 GCTGCATATCAGATCATAAAGATGACGTGATGCTTGGTAACTCCCGAGGCTTAACCTA 193
Qy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
Db 194 GCTTCGTTTTGACTACATGATGAAACCGAGAGTGTGACAAACTCATATGGAATCTATC 253
Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
Db 254 AACAAAGAACTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 313
Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
Db 314 GTAACATATAGCTCCACTGTTCAATGCCCCACTCCAGGNAATCTGAGACGGCGGTGGGA 373
Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
Db 374 GTAGGACAGTGGTTCTTCCAGAAAGCCATCATGTAGCCGGATTTGGCTTCAAAAAGAAA 433
Qy 141 TrpGlnAsnLysArgLysAlaGluGlySerProValAspLysProAsnIleValThrGly 160
Db 434 TGGCAGAAACAACCGCAGAGCTGAGGGTAAACCCCTATGACAAACCCCAACATTTGTCAC 493
Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
Db 494 GCCAAATGTTCAAGTTTCTGGGAGAAATTCGCTCGGTACTTCGAGGTGGAGCTAAAGAA 553
Qy 181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
Db 554 GTAAACCTAAGTGAAGTTACTACGTGATGGATCCAGACAAAGCAGCAGAAATGGTAGAC 613
Qy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrIleAsnGlyGluPheGluAsp 220
Db 614 GAGAACACAATCTGTGTCGACCATATTTGGGATCCACACTCAACCGGTGAGTTCCGAAGAC 673
Qy 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrPheThrPro 240

```

```

181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
541 GTCAACTAAGAGAGACTATTACGTGATGGACCTGTAAAGCGCGTCAAAATGCTAGAC 600
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
601 GAAACACAATTTGTCGCTCCATCTCCGTTCAACCGTTAAACCGGTGAATTCGAAGAC 660
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
661 GTTAAAGCTCTCAACGACCTCTTGTGAGAAAACAAGCAAAACCGGATGGACACGCCA 720
241 IleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
721 ATACAGTGGACGACGAGTGGTGGTTAATGCTCCGTTCTTGTATCCGGAGCTGGAG 780
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
781 TGGGATTTCCGGCTACCGTTGGTAAAGATTAATAATGTGAGTGGTCAAAAATACGGTTG 840
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeu 300
841 GTTACGCCGCTATGTTGGTGGTGTATGGAGAACCAACCGAATTCCTCATGACTT 900
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
901 ATCTTCCATATCAATTAATCTTGGCGTGTATCAACCAACCTTTACACTCAACTTCCAAA 960
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
961 GGTTCAAGTCAAGTGAATGCTCAGTACTACCAGCTGATTCGCTTGGATTCGAGGTTAT 1020
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
1021 CGCAATGATGATGATTAATGTCGGGAAACAATGATGATGATGATGATGATGATGATGAT 1080
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
1081 ACGGGACGTTTTAAATTCGTTCCAAAGAAAACGGTGTTCGTTAGTGGCGTTTTCTCTC 1140
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTyr 400
1141 AAGATAGTAGCCGCCACACAGAGTTCGAGGTGGCCCAATACACTCGCTTCGCTCGGCTGG 1200
401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
1201 ATCGTTCCGGCTACAGATGCTCGGATGCGGATGCGCAGCAATGCTACTCTCGAGTTGT 1260
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
1261 ATCCGAGAAGATTTCTCTCGAACTTAGCCGAGAGATGGTACTGATTTCCGAGAAGGTT 1320
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1321 CTACAGAGCTCGATACCTTCCGGCGAGGTTACGCCAGATGCTAATGGAAAAGTT 1380
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1381 AACCGT-----GTTAAGAAAGCCGCCAGAGGAGACG 1410
481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysIleThrSer 499
1411 CAGAGAAAGTCAAGCTACTCGAAGAAGTTGTTG---GAGACTAAGAGACCAAC 1464
RESULT 9
US-10-006-852-3
Sequence 3, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
APPLICANT: Kinnersley, Alan M.
APPLICANT: Turano, Frank J.
TITLE OF INVENTION: Methods for Regulating Plant GABA Production

```

```

674 GTGAACCGTCTCAATCACTGCTAGTCAAGAAAAACGAGGAGACTGGTTGGAACACACCG 733
241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
734 ATCCACGTTGGATGACAGAAAGTGGAGGTTTCAAGTCCCGTTTATCTATCCCTGAATGAA 793
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
794 TGGGACTTTAGACTTCTTCTGGTTAAGATATCAACCGTGGTGGTCAACAGTATGACTG 853
281 ValTyrAlaGlyIleGlyTyrValIleTyrArgAsnLysGluAspLeuProGluLeu 300
854 GTCATATGCTGGTATGGTGGTGGTGGGAGGAGGAGGATTCCTGAAAGGCTT 913
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
914 AICTTTCAATTAATTAATCTTGGTGTGATCAACCCACTTTCACCTCAATTTCTCBAAG 973
321 GlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
974 GGATCGAGCCAAATTAATGCTCAATACTACAGCTCATTCGCTTGGATTCGAGGGGTAC 1033
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluLeuGluLys 360
1034 AAAAATGATGGAGAAATGCAATAGAGAAATCGTGGTGGTTCCTAAGAGAGGATAGAGAA 1093
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
1094 ACAGAGCGTTTCAACATAGTCTCAAGAGGACCAAGGAGTCCAGTCCGATGCTCTCTC 1153
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTyr 400
1154 AAGGACCATAGTTCCACAAACGAGTTCGAGATCTCTGAGATCTACGTCGTTTGGCTGG 1213
401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
1214 ATCGTCCCAGCTTACACTATGCTCCGATGACAGCACATCACGGTTCGCGTGTGTC 1273
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
1274 ATCAGGAGAGATTTCTCAAGAACACTCGCGGAGACTTGTGCTGATATTTCGAAGGTG 1333
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1334 CTTTCATGAGTAGATACCTTTCCTTCCAAAGATATCTAAGAAGATGGGAATA----- 1384
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1385 -----GAAGGATCGCGGAAATGTAAAGAGAGAAAGATGGAGAG 1426
481 Gln-----ArgAspIleIleThrGlyTyrLysPheValAlaAspArgLysThr 498
1427 GAGATCTGTATGGAAGTATTGTTGGATGGAGGAACTTCTGAAAGGAGAGGAAAGATG 1486
499 SerGlyIleCys 502
1487 AATGGTGTGTGC 1498

```

```

RESULT 10
S-10-167-547C-1
Sequence 1, Application US/10167547C
Publication No. US20030170653A1
GENERAL INFORMATION:
APPLICANT: E.I. du Pont de Nemours and Company
APPLICANT: Damude, Howard G.
TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
BUTYROLACTONE AND ITS INTERMEDIATES
FILE REFERENCE: CUI804 US NA
CURRENT APPLICATION NUMBER: US/10/167,547C
PRIORITY FILING DATE: 2003-03-17
PRIOR APPLICATION NUMBER: 60/297198
PRIOR FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 67

```

```

; SOFTWARE: Microsoft Office 07
; SEQ ID NO 1
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Alstroemeria
US-10-167-547C-1
Alignment Scores:
Pred. No.: 1,44e-231 Length: 1497
Score: 2073.00 Matches: 391
Percent Similarity: 87.92% Conservative: 53
Best Local Similarity: 77.43% Mismatches: 51
Query Match: 79.27% Indels: 10
DB: 12 Gaps: 2
US-10-006-852-2 (1-502) x US-10-167-547C-1 (1-1497)
Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 1 ATGTTCTCTCCAGCGCGCTCCGACACACACCGGCCAGTCCACCTGCACCTTCGGGTG 60
Qy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 61 CGTACGTCGCGATGCGCCTGCCGGTTCAGGATGCCGGAGATTCGATACCCAGGAC 120
Qy 41 AlaAlaTyrGlnIleIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 121 ACGCGTACAGATGCTCAACAGCAGCTGATGCTGGACGGGAAACCCAGGCTGAACCTG 180
Qy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerIle 80
Db 181 GCGTCGTTGTCACGACTCGATGGAGCCCGAGTGGATCGCTCATGATCGCCGCC 240
Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
Db 241 AACAAAGTATGTCGACATGAGCAGTACCCGGTCAACAACCGAGCTACAGAATCGCTGT 300
Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
Db 301 GTTAATAATAGTCCACTTTCATGCGCTTATGGGGATGAGGAAACAGCAGTAGGA 360
Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
Db 361 GTTGGACAGTGGGGTTCATCAGAACCAATATGCTCAGCTTGGCTTGGCTTCAAGAGAG 420
Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
Db 421 TGGCAGAACAAAAAAGCAGAGGGGAAAGCCATATGACAAAGCCCAATATGTCAACCG 480
Qy 161 AlaAsnValGlnValCysTyrGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
Db 481 GCAACCGTTCAGGTTCTGGGAGAAATTCCTAGGATTTTGAAGTTGAACTGAAAGAA 540
Qy 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnAlaValAspMetValAsp 200
Db 541 GTGAAGCTGAGGGAGGTTTATACATCATGGACCACAGAGGCTGTGGAAATGGTGGAT 600
Qy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
Db 601 GAGAAATACCATATGTTGCTGCTATCTTGGCTCAACCCCTTACTGGAGAGTTCGAAAG 660
Qy 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrPro 240
Db 661 GTTAACCTACTGAAACAACTTCTTGAAGAGAAACAAAGAAACTGGTGGGACACACCC 720
Qy 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
Db 721 ATTCAATGTTGATGCTGCTAGTGGTGGATTCATGCTCTCTTCTATACCCAGAACTG 780
Qy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
Db 781 TGGGATTTCCGATTTACCACCTGGTGAAGATTAATAATGTCAGCGGACACAAATATGG 840

```

281 VAlTrAlaGlylleclYtrpValIletrpArgAsnLysGluAspLeuProGluLeu 300
 841 GTTATTCAGGGTCGGTTGGTGTCTGGAGGAAACAAGAGATCTTCTTGAAGCTC 900
 301 IlePheHisIleAsnTyrluGlyAlaAspGlnProThrPheThrIleuAsnPheSerLys 320
 901 ATTTTCCATATAACTACCCTTTGGGGGAGATPCAGCCCTACTTTCCACCCTCAATTTCTCTAAA 960
 321 GlyserserGlnValIleAlaGlnTyrTrpGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 961 GGTCAAGCCAGATAATTGCTCAATTATTAATTAATTCATTCCGCTTGGTTTTCCAGGGGTAT 1020
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 1021 AAGAACAATAATGTAATAACTGCATGAGGAGAACACAAGAATACTGAGAGAGACTCTGCAGGAG 1080
 361 ThrGluArgPheAsnIleValSerLysAspGluValProLeuValAlaPheSerLeu 380
 1081 ACGGGCCGTTTCGAGATAGTCTCAAGAATATTGGGGTGCCTCTTGTTCATTGCTCTC 1140
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
 1141 AAGCACAGCAGTCAGTACACTCTTTGAGATAGCGAGCAGCCATGAGAGAGTTCGGATGG 1200
 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1201 ATCAATTCCTGCATACCAATAGTCCCAAGGACCGGAGCACATAGCTGTCCCTCGTGTGTT 1260
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1261 ATCAGGGAGATTCAGCAGGAGCTTCTGTCAGCGCCTAGTTAATGACATGAAAGAGGTG 1320
 441 MetArgGluLeuAspGluLeuProSerArg-----ValIleHisLysIleSerLeu 457
 1321 CTGGCTGAGCTGGACCTACTTCCAGTCCGATCACCACCATTGCCCATCTWACGGCTGTG 1380
 458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAsp 477
 1381 -----GAGAAGATATAAGCGAAGCTGTGATCAAGAGAGTTTC 1419
 478 IleAspLysGlnArgAspIleIleThrGlyTrpLysPheValAlaAspArgLysLys 497
 1420 CTCGAGATAGAGAAGAGTATTACACATTGGAAGGATAGTGTGATGAAACCGCGAGAG 1479
 498 ThrSerGlyIleCys 502
 1480 ACTAATAAAGTTTTC 1494

ESULT 11
 S-10-006-852-17
 Sequence 17, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
 APPLICANT: Kinnersley, Alan M.
 APPLICANT: Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 17
 LENGTH: 1783
 TYPE: DNA
 ORGANISM: Lycopersicon esculentum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (6)..(1511)
 OTHER INFORMATION:
 US-10-006-852-17

Alignment Scores: 1.46e-224 Length: 1783
 Pred. No.: 2014.00 Matches: 387
 Score: 86.44% Conservative: 53
 Percent Similarity: 76.03% Mismatch: 55
 Best Local Similarity: 76.03% Indels: 14
 Query Match: 17.02% Gaps: 5
 DB: 14

US-10-006-852-2 (1-502) x US-10-006-852-17 (1-1783)

Qy 1 MetValLeu--SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19
 Db ATGGTGTAAACAACCGATCAAGAGATTCAAGAGAGCTTGCACCTGTCACATTTGCA 65
 Qy 20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
 Db TCAAGATATGACAGCACTTACCTTAGTTCACAAAATGCCCTAAATAAAATCCATGCCGAAA 125
 Qy 40 GluAlaIleTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
 Db GAAGCAGCTTATCAGATTGTAACACGACGAGCTTATGTTGGATTGTAACCCAGGTTGAAT 185
 Qy 60 LeuAlaSerPheValThrTrpMetGluProGluCysAspLysIleIleMetSerSer 79
 Db TTAGCTTCTTTTGTAGCATGATGGAGCCCGAGTGCATGATAGCTCACTGATGATCTCATCC 245
 Qy 80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
 Db ATTAATAAATAAATGTCGACATGGATGAGTATCTTCCACCACTGAATTCATAAATAGA 305
 Qy 100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
 Db TGTGTTAACATGTAGCACATCTTTCCATGCGCCCGTGTGATGATGAGACTGCAAGTT 365
 Qy 120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
 Db GGGTTGCTACAGTGGTTCATCAGAGCCATAAATGCTTGTGSCCTTGTCTTTCACAACCC 425
 Qy 140 LysTrpGlnAsnLysArgLysAlaGlyLysProValAspLysProAsnIleValThr 159
 Db AATATGCATTCGAAAAGAAAGAGGACCTTTCGATAAAGCCATAATATAGTCACT 485
 Qy 160 GlyValAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
 Db GGAGCTAATGTGAGGTCCTGCGAAATAATTCAGAGGTATTTTTCAGGTTGAGTTGAA 545
 Qy 180 GluValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetVal 199
 Db GAGTGAATACTAAAAGAGGATACTATGTAATGGACCTGCCAAAGCAGTAGAGATAGT 605
 Qy 200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
 Db GATGAGAATAACAATATGTGTTGCTGCAATCTTGGTTTACTCTGACTGGGAGTTTGA 665
 Qy 220 AspValLysLeuLeuAsnAspLeuLeuValGlyLysAsnLysLysGluThrGlyTrpAspThr 239
 Db GATGTAAGTCTCTAAACGAGCTCCCTTCAAAAAPAAACAGGAAACCAGATGCGAGACA 725
 Qy 240 ProIleHisValAspAlaIleAsnSerGlyPheIleAlaProPheLeuTyrProGluLeu 259
 Db CCGATTCATGCTGCTGGAGTGCAGTGGAGGATTTATTTGCTTTCTCTTCCCTGCCAGATCTT 785
 Qy 260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
 Db GAAATGGGATTTCCGTTGGCCCTCTTTGTAAGATATAAATGTGAGCGGTCCAAAGTATGCC 845
 Qy 280 LeuValTyrAlaGlyIleGlyTrpValIleTyrArgAsnLysGluAspLeuProGluGlu 299
 Db CTGATATGCTGGTGTGGTGGGTGATATGCGGGAGCAAGAGAGACTTGGCCCGATGAA 905
 Qy 300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
 Db CTGCTCTTTTCATATAAATACTACCTTGGGTCTGATGATCAGCTTATCTTCTCAACTTCTCT 965

320 LysGlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGlyHisGluGly 339
 966 AAAGTTCTCTATCAAATAATGACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGT 1025
 340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
 1026 TATAAGAACGTCATGAGAATTGCTTTATCAAAACGCAAAAGTACTAACAAGAGGAATCACA 1085
 360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
 1086 AAAAAATGGGGGGTTCGATATTGCTCTAAGGATGCGGTTCCTGTTGTAGCAATTTCT 1145
 380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGly 399
 1146 CTCAGGCAGCAGCAAAATATACGGTATTGAAATATCTCAGCAATCTCAGAAGAATTTGA 1205
 400 TrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgVal 419
 1206 TGGATCGTCCCTCAVACAAATGCCCACCGGATGCTGAACACAAATCTCTACTTGGGGTT 1265
 420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
 1266 GTCATTAGAGAGGATTCAGCCACAGCTAGCTGAGAGACTTGTTCGACATTTGAGAAA 1325
 440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle-----SerIleu 457
 1326 ATTCCTTCAGAGTGGACACACACAGCCTCTCGTTTGGCCCAAAAGCTCTCGGTCTCACT 1385
 458 GlyGlnGlySerGluSerAsnSerAspAsnLeu-----MetValThrVal 473
 1386 GCTGAGAGAGTGGTATGACAAAGGTGATGGGCTTCATCAATTTTCACATGGATCTGTA 1445
 474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAla 493
 1446 -----GAGACTCAGAAAAGCATTATCAAAGATTGGAGGAAAATCGCAGGG 1490
 494 AspArgLysLysThrSerGlyIleCys 502
 1491 -----AAGAAGACCAGCGGAGTCTGC 1511

RESULT 12
 S-09-887-576-782
 Sequence 782, Application US/09887576
 Patent No. US20020144047A1
 GENERAL INFORMATION:
 APPLICANT: Bugworth, P.
 APPLICANT: Brown, D.
 APPLICANT: Chang, H.
 APPLICANT: Zhu, T.
 APPLICANT: Han, B.
 APPLICANT: Wang, X.
 APPLICANT: Cooper, Bret
 TITLE OF INVENTION: Promoters for regulation of plant expression
 FILE REFERENCE: 1360.001US1
 CURRENT APPLICATION NUMBER: US/09/887,576
 PRIOR FILING DATE: 2001-06-25
 PRIOR APPLICATION NUMBER: US 60/213,848
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 60/214,087
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 60/258,692
 PRIOR FILING DATE: 2000-12-29
 NUMBER OF SEQ ID NOS: 875
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 782
 LENGTH: 1455
 TYPE: DNA
 ORGANISM: Oryza sativa
 JS-09-887-576-782
 Alignment Scores:
 Pred. No.: 6.97e-224 Length: 1455

Score: 2007.00 Matches: 381
 Percent Similarity: 87.67% Conservative: 60
 Best Local Similarity: 75.75% Mismatches: 42
 Query Match: 76.75% Indels: 20
 DB: 10 Gaps: 6
 US-10-006-852-2 (1-502) x US-09-887-576-782 (1-1455)
 QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHis---SerThrPheAla 19
 Db 1 ATGGCGCTGTTCGACGGCG-----CAGACAGGGGAGTGCATGCTCCTCGACCTCGACGTCGCG 54
 QY 20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
 Db 55 TCGCGGTACGTCGCGACGGCTCCGAGGFTTCAGGATGCCGAGAAATCGATCCCCAAG 114
 QY 40 GluAlaIaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
 Db 115 GACGGGGGTACCAGATCATCAACGACGAGGTGATGCTCGACGGCAACCCCGCGCTGAAC 174
 QY 60 LeuAlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
 Db 175 CTGGGCTCTTCTGTCACCCAGTGGATGGAGCCGAGTCCGACAAAGCTCATGATGGCGCC 234
 QY 80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
 Db 235 ATCAACAAGAACTACGTCGACATGATGATGATACCCCGTCAACCGAGCTCCAGAACCGG 294
 QY 100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
 Db 295 TGCCTGAACATGATCGCGCATCTGTTCAACGCGCGATCGGGGACGACGAGACCGCGCTC 354
 QY 120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
 Db 355 GGGGTGGGCACGGTGGGTCTCGAGGCCATCATGCTGGCGGGCTGGCGTTCAAGAGG 414
 QY 140 LysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThr 159
 Db 415 AAGTGCAGAACAGGATGAAGCCGAGGGGAAAGCCCCACGACAGCCCAACATCTGTGAG 474
 QY 160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
 Db 475 GGGGCCAAGTGCAGTGTGCTGGGAGAAAGTTCGGCGCTACTTTCGAGGTGGAGTCAAG 534
 QY 180 GluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetVal 199
 Db 535 GAGGTGAAAGTGCACCAAGGGTACTACGTCGATGTAACCCGCGAGAAAGCCGCGGATGTC 594
 QY 200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
 Db 595 GACGAGAACACCATCTGGCTCGCCCAATCTCGGCTCCACCTCAACGGCGAGTTCGAG 654
 QY 220 AspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
 Db 655 GACGTCAGATGCTCAACGACCTCTCCAGCCAAAGAAAGCCGCGAGACAGGTTGGAACAG 714
 QY 240 ProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
 Db 715 CCGATCCATGTGGACGGCGGAGCGGGGTTTCATCGCGCGCTTCATACCCCGAGGTG 774
 QY 260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
 Db 775 GAGTGGGACTTCCCGCTCCCGTGGTGAAGAGATCAACGTCAGCGGGCCACAAAGTACCGG 834
 QY 280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
 Db 835 CTCGTCAGCCCGGGCTCGGGTGGTTCATCTGGCGCAACAAGAGGACCTCCCGGATGAG 894
 QY 300 LeuIlePheHisIleAsnTyrLeuGlyAlaAlaAspGlnProThrPheThrLeuAsnPheSer 319
 Db 895 CTCATCTTCCACATCAACTACTCCTCGCGCCGACAGCCAACTTCCAGCTCAACTTCTCC 954
 QY 320 LysGlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGlyHisGluGly 339

```

> 955 AAAGGATCGAACAGATAATTGGCCAGTATTACAGCTCATTGGTCTGGATTCCGAGGGG 1014
/ 340 TTTArgAsnValMetGluAsnCysArgGluAsnMetIleValIleuArgGluGlyLeuGlu 359
> 1015 TCAAGGACATCATGACAGAACTCCCGGGCAACGCGACCGGGTCTCCGGGAGGGATPCGAG 1074
/ 360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProIleuValAlaPheSer 379
> 1075 AAGACGGCCACTTCGACGTGGTCCAAAGACTCCGGGCTGCCGCTGGTGGCCCTTC 1134
/ 380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
1135 CTCAAAGACTCGTCGGGTACACGGTGTTCGAGGTGGCCGAGAGCCTCCGGCCGCTTCGGC 1194
/ 400 TrpIleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgVal 419
> 1195 TGGATGTGCCGGCGTACCAATGCCCGCGGACCTGACGACGTCCCGGTGATGGCGGTC 1254
/ 420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
> 1255 GTCATCCGGGAGGACTTCAGCCCGCGCTCCCGAGCGCTCATCACCGACTCACCAG 1314
/ 440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisIysIleSerLeuGlyIle 459
> 1315 ACGGTGCCGATATGGAC-----GCCCCACCGCGTC----- 1344
/ 460 GluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAsp 479
> 1345 AAGAAGCCGCCCGCCGAG-----CCGGCCAAAGAACCCGTGGCGGGAG 1386
/ 480 LysGlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysThrSer 499
> 1387 ATAGAGAAGAGGTGACCACTACTCGCGGAGTTTCGTCCGCG---AGGAAGAAGACGAC 1443
/ 500 GlyIleCys 502
b 1444 CTCGTCTGC 1452

```

ESULT 13

S-10-006-852-7

Sequence 7, Application US/10006852

Publication No. US20030046732A1

GENERAL INFORMATION:

APPLICANT: Kimmersey, Alan M.

TITLE OF INVENTION: Methods for Regulating Plant GABA Production

FILE REFERENCE: 7224-65

CURRENT APPLICATION NUMBER: US/10/006,852

PRIOR FILING DATE: 2002-07-01

PRIOR APPLICATION NUMBER: US 60/246,367

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.1

SEQ ID NO 7

LENGTH: 2121

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(87)

OTHER INFORMATION:

FEATURE:

NAME/KEY: CDS

LOCATION: (274)..(681)

OTHER INFORMATION:

FEATURE:

NAME/KEY: CDS

LOCATION: (782)..(994)

OTHER INFORMATION:

FEATURE:

NAME/KEY: CDS

LOCATION: (1081)..(1335)

```

/ OTHER INFORMATION:
/ FEATURE: CDS
/ NAME/KEY: CDS
/ LOCATION: (1438)..(1488)
/ OTHER INFORMATION:
/ FEATURE: CDS
/ NAME/KEY: CDS
/ LOCATION: (1654)..(2118)
/ OTHER INFORMATION:
US-10-006-852-7

```

```

Alignment Scores: 2,13e-215 Length: 2121
Pred. No.: 1936.50 Matches: 410
Score: 61.99% Conservative: 32
Percent Similarity: 57.50% Mismatches: 46
Best Local Similarity: 74.05% Indels: 226
Query Match: 14 Gaps: 7
DB:

```

```

US-10-006-852-2 (1-502) x US-10-006-852-7 (1-2121)
QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 1 ATGGTTTTCGCTAAGACAGTTCCGAACTCGAATCTCAATCTCAATCCAACTTTTGGCTT 60
QY 21 ArgTyrValArgThrSerLeuPro----- 28
Db 61 CGTTACCTCCGCAACTCTCTCCACGGTAACAACACTTGTAAACACAAAATCTTTTGTCTAATG 120
QY 28 ----- 28
Db 121 TTTTCGTCAACAATAGTAAACAATGATGATAAACCCTTGGATAGTTTTTTTTTTGGCCG 180
QY 28 ----- 28
Db 181 TGGTTAATGTTGTAGATTATTATTATGTTTATATACTATAAAGGAAGGACATGTTTCGTATT 240
QY 29 -----ArgPheLysMetProGluAsnSerIlePro 38
Db 241 TTTAACTTAATGATCATCACTTTTCATCTAGATAGAAATCGAAATGCTGAGAACTCAATCCCA 300
QY 39 LysGluAlaIaIaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeu 58
Db 301 AAGAAGCAGCTTACCRAATCAICACGACGAGCTAATGCTCGATGCTAACCCACGAGCTG 360
QY 59 AsnLeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSer 78
Db 361 AACCTAGCTTCTTCGTGACCAATGATGATGGAGCCAGAAATGTGACAAAGCTCATGATGGAG 420
QY 79 SerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsn 98
Db 421 TCCATCAACAAAGAACTACGTCGACATGGAGTACCCTGTCCACCTGAGCTTACAGAAC 480
QY 99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAla 118
Db 481 CGATGTGTTAAATGATAGCAGCTCTTTCACCGCGCGCTTGGTGACGGTGAAGCTGCC 540
QY 119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
Db 541 GTTCGTGTTGGCACCCGTCGGATCGTCGGAGCGAATATTGTTGGCCGGTTTGGCTTTAAG 600
QY 139 ArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleVal 158
Db 601 AGCAATGGCAGATTAAGCCGTAAGGCCCAAGGCTTCTTATTGATTAAGCCCAATATCTGTA 660
QY 159 ThrGlyAlaAsnVal----- 163
Db 661 ACCGGTCTAATGTCCAGGTAAACAAAACAAAATTGATGAAATATTAACCAAGACAAA 720
QY 164 -----GlnVal 165
Db 721 ATTGAATTTTCAATCCGGTTAAGTTATATATGTTGACTCAATTTCCGGTTCAATACAGGT 780

```

165 lCysTrpGluLeuPheAlaArgTyrPheGluValGluLeuLysGluValIleLeuSerG1 185
 781 TTCTGGAGAAAATCCCAAGGTATTTCGAAGTGGAGCTTAAGAAAGTGAACCTAAGAGA 840
 185 uG1yTyTyrValMetAspProGlnGlnAlaValAspMetValAsnThrlleCy 205
 841 AGACTATTAAGTATGATGACCCCTGTAAGGGGGTCAAAATGGTAGACGAAACAAATTTG 900
 205 sValAlaAspIleLeuGlySerThrLeuAsnGlyLupheGluAspValLysLeuLeuAs 225
 901 TGTCGCTGCATCCTCGGTTCAACGTTCACCGGTTCAATTCRAAGAGCTTAAGCTTCCTCA 960
 225 nAspLeuValGluLysAsnLysGluThrGly 236
 961 CGACCTCTTTGTCAGAAAAACAAGCAAAACCGG-GTAAITTAACCAACCGAGAAAACAAG 1019
 236 236 ----- 236
 1020 CTAATATFCGATTTGTAATTCGGTTGGAGTCCGGTTTAAACGTTTAAAACAAATTTGCAG 1079
 237 -TrpAspThrProIleHisValAspAlaIleAspGlyGlyPheIleAlaProPheLeuTy 256
 1080 ATGGGACACCCCAATACACGTGACGAGCGAGTGGTGGTTTATTCCTCCGTCTTTGTA 1139
 256 rProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHi 276
 1140 TCCGGAGCTGGAGTGGAGTTCGGCTTCCGGTTGGTTTAAAGAGTATTAAATGTAGAGTGTCA 1199
 276 sLysTyGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysLusPlE 296
 1200 CAATACGGTTTGGTTTACCGCCGGTATGTTGGTTGTAGAGAACCAAAACCGATTT 1259
 296 uProGluLeuIlePheHisIleAsnTyLeuGlyValAspGlnProThrPheThrlE 316
 1260 GCCTGATGAATTCCTCCATCAATTCATTCGCGCTGTACCAACCAACCTTTACACT 1319
 316 uAsnPheSerLys 320
 1320 CAACCTTCACAAAGGTAATTACCATAAGTCCATAATATACTTTCAATTAATATTT 1379
 321 -----GlySe 322
 1380 TTGGTGTATGGAATGTTTATAGACTAAACATTTGATAATGTTGTATAAACCAGGTTC 1439
 322 rSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGlu 338
 1440 AAGTCAAGTATGCTCAGTACTACCAGCTGATTCGCTCTTGGATTTCGAGTAAATAATA 1499
 338 ----- 338
 1500 CTCAAAATAGCAATATAATTTACCATAAGTGGTCAATAAAGAAACTAGAAATGATTTATTTA 1559
 338 ----- 338
 1560 AGTGTACTTGTTACTACTTGAATTAACCTTCCCAACGTTCCTAACATGACTAGTTTTGGTATT 1619
 339 -----GlyTyrArgAsnValMetGluAsnCy 347
 1620 GTGTAAATTAATAATGTTTTCCTGTTTGTATAGGGTTATCCGAATGTGTAGTATAATTG 1679
 347 sArgGluAsnMetIleValLeuArgGluGlyLeuGlyLysThrGluArgPheAsnIleVa 367
 1680 TCGGAAACAACATGATGTACTAAGCAAGGATTAAGAAACCGGACGTTTTAARAATFCGT 1739
 367 lSerIleAspGluGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHisTh 387
 1740 CTCAAAAGAAACCGTGTTCCTGTTAGTGGCTTTTCTCTCAGAAGATAGTACCCCAACAA 1799
 387 rGluPheGluIleSerAspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMe 407
 1800 CGAGTTCGAGGTGGCCATACACTCCGTCGCTTCGGCTGGATCGTTCGGGCCITACAGAT 1859
 407 tProPheAsnAlaGlnHisIleThrValLeuArgValIleArgGluAspPheSerAr 427

1860 GCCTCCGATCGCAGCATGTCACACTGCTCCTTCGAGTTGTTATCCGAGAAAGATTTCTCTCG 1919
 427 gThrLeuAlaGluArgLeuValIleAspIleGluLysValMetArgGluLeuAspGluLe 447
 1920 AACCTTACCCAGAGATTGGTAGCTGATTTCGAGAAGTTCACACGAGCTTCGATACCGCT 1979
 447 uProSerArgValIleHisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSerAs 467
 1980 TCCGGCGAGGTTCACGCCAAGATGCTTAAATGGAAGAGTTAACCGT----- 2025
 467 pAsnLeuMetValThrValLysLysSerAspIleAspLysGlnArgAspIleIleThrG1 487
 2026 -----GTTAAAGAGACCCGAGGAGACGAGGAGAGGAGGAGAGAGTAAAGTCAACCGCCCTA 2069
 487 yTriPlysLysPheValAlaAspArgLysLysThrSer 499
 2070 CTGGGAAAGTGGTTG---GAGACTAAGGAGACCAC 2103

RESULT 14
 US-09-887-576-806
 ; Sequence 806, Application US/098887576
 ; Patent No. US20020144047A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Budworth, P.
 ; APPLICANT: Brown, D.
 ; APPLICANT: Chang, H.
 ; APPLICANT: Zhu, T.
 ; APPLICANT: Han, B.
 ; APPLICANT: Wang, X.
 ; APPLICANT: Cooper, Bret
 ; FILE OF INVENTION: Promoters for regulation of plant expression
 ; FILE REFERENCE: 1360.001US1
 ; CURRENT APPLICATION NUMBER: US/09/887,576
 ; PRIOR FILING DATE: 2001-06-25
 ; PRIOR APPLICATION NUMBER: US 60/213,848
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 60/214,087
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 60/258,692
 ; PRIOR FILING DATE: 2000-12-29
 ; NUMBER OF SEQ ID NOS: 875
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 806
 ; LENGTH: 1287
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1287)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-887-576-806

Alignment Scores:
 Pred. No.: 2,97e-210 Length: 1287
 Score: 1889.50 Matches: 362
 Percent Similarity: 91.47% Conservative: 24
 Best Local Similarity: 85.78% Mismatches: 20
 Query Match: 72.26% Indels: 17
 DB: 10 Gaps: 1

US-10-006-852-2 (1-502) x US-09-887-576-806 (1-1287)

QY 32 MetProGluAsnSerIleProLysGluAlaIleAlaTyrGlnIleIleAsnAspGluLeuMet 51
 Db 1 ATGCCGGAGACTGCTTCCCAAGGGGGGGGTACCAGATCATCACGACGAGCTGATG 60
 QY 52 LeuAspGlyAsnProArgLeuAsnLeuAlaSerPheValThrTrpMetGluProGlu 71
 Db 61 CTGGACGGCAACCCCGGCTGAACCTCGCTCGTTCGTCCACCGCTGGATGGAGCCCGAG 120
 QY 72 CysAspLysLeuIleMetSerSerIleAsnLysAsnTyValAspMetAspGluTyPro 91

121 TGGCACAAGCTCATCCAGGCTCCGTCACAAAGAACTACGTCGACATGCGACGAGTACCCC 180
 92 ValThrThrGluLeuGlnAsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaPro 111
 181 GTCCACCGAAGCTCCAGG-----
 112 LeuGluGluAlaGluThrAlaValGly-ValGlyThrValGlySerSerGluAlaIleMe 131
 201 -----GACTCTGAACCGGCTCGGAAGTGGCACTGTCCGGCTCGCTCGAGCCATCAT 254
 131 tLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLysArgLysAlaGluLysPr 151
 255 GCTCCGGGTTGGCTTCAAGAGGAGTGGCAGAACAAAGATGAAGCAGCCGGCAAGCC 314
 151 oValAspLysProAsnIleValThrGlyAlaAsnValGlnValCysTrpGluLysPheAl 171
 315 ATGGCACAAGCCTAACATTCTCCAGCGCCCAATGTCCTCAAGTTTGTGGGAGAGTTCCG 374
 171 aArgTyrPheGluValGluLeuLysGluValLysLeuSerGluGlyTyrValMetAs 191
 375 GCGATATCTCGAGGTTGAGCTCAAGAAAGTGAAGCTGAGTGCAGCGGTACTACTGTCATGGA 434
 191 pProGlnGlnAlaValAspMetValAspGluAsnThrIleCysValAlaAspIleLeuG 211
 435 CCACGCTPAGCCGTGGATATGTCGACGAGAACACCACTCCGCTCCGGCGATCTCCGG 494
 211 YSerThrLeuAsnGlyGluPheGluAspValLysLeuLeuAsnAspLeuLeuValGlu 231
 495 GTCCAGCCTGAACCGGGAGTTCGAGGACGTGAAGCTGCTCAACGATCTGCTCACCAAGAA 554
 231 sAsnLysGluThrGlyTrpAspThrProIleHisValAspAlaAsnSerGlyPheI 251
 555 GAACGCTGAACAGGCTGGACACCGCATCCACCTGGACCGCGGAGCGGGGTTTCAT 614
 251 eAlaProPheLeuTyrProGluLeuGluTrpAspPheArgLeuProLeuValLysSerI 271
 615 CGGCGGTTCTGTACCGGAGCTGGAGTGGACTTCCGGCTCCCGCTGGTGAAGACAT 674
 271 eAsnValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpAr 291
 675 CAACGCTGAGCGGGCACAAAGTACGGCCCTCGTCTACCGCGGATCGGCTGTCATCGGAG 734
 291 sAsnLysGluAspLeuProGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspG 311
 735 GAGCAGGAGGATCTGCTGAGGAGCTCATCTTCCACATCAACTACTCCGCGCGCACCA 794
 311 nProThrPheThrLeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrGly 331
 795 GCCCACCTTACCCTCAACTTCCCAAGGTTCCAGCCAGTCTATTCACAGTATTACCA 854
 331 nLeuIleArgLeuGlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMe 351
 855 ACTAATCCGCTTAGGCTTTGAGGGGTACAAGAACATCATGGAGAACTCCAGGAGAACCG 914
 351 tIleValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspG 371
 915 GATGGTCTGAGCAGGGCTGGAGAGAGCGGGCGGTTCAACATCGTGTGCCAAGGACAA 974
 371 uGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluI 391
 975 CGGGTCCGGCTGGTGGCTTCTCCCTCAGGACAGCGCCCGGCAACACCGAGTTCGAGAT 1034
 391 eSerAspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProAsnAl 411
 1035 CTCCGACTTCTCCGCGCTTCCGCTGGATCGTCCGCGCTTACACCATGCGCCCGCCGACGC 1094
 411 aGlnHisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaG 431
 1095 CGACAGCTCACCGTCTCCCGTCTCCGCTGCTCCGCGAGACTTTCAGCCGACCGCTCGCCGA 1154
 431 uArgLeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgVa 451
 1155 GCGCCTCGTCTCGAGCTCGAGAGAGGTTCTGCACGAGCTTCGACGCGCTTCCCGCGCCGCGT 1214

451 lile 452
 1215 CGTC 1218
 RESULT 15
 US-10-167-547C-3
 ; Sequence 3, Application US/10167547C
 ; Publication No. US20030170653A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E. I. du Pont de Nemours and Company
 ; APPLICANT: Damude, Howard G.
 ; TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
 ; FILE REFERENCE: CL1804 US NA
 ; CURRENT APPLICATION NUMBER: US/10/167,547C
 ; CURRENT FILING DATE: 2003-03-17
 ; PRIOR APPLICATION NUMBER: 60/297198
 ; PRIOR FILING DATE: 2001-06-08
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: Microsoft Office 07
 ; SEQ ID NO 3
 ; LENGTH: 1530
 ; TYPE: DNA
 ; ORGANISM: Alstroemeria
 US-10-167-547C-3
 Alignment Scores:
 Pred. No.: 3,698-207 Length: 1530
 Score: 1864.00 Matches: 351
 Percent Similarity: 84.55% Conservative: 54
 Best Local Similarity: 73.28% Mismatches: 70
 Query Match: 71.28% Indels: 4
 Gaps: 2
 US-10-006-852-2 (1-502) x US-10-167-547C-3 (1-1530)
 QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 1 ATGGCTCTCCAGCGCTGCTCCGACTCCAAACCAACCAAGTGCAGTGCACCTATGCCTCT 60
 QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 61 CGCTACGTTTCGGCAGCAGGCTCCAGGTTTCAGGATCCCGGAGAGTTCGATACCAAGAG 120
 QY 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 Db 121 GCGGGTTCACGATGATCAACGACGAGCTGATGCTGGACGGGAACCCCGAGCTGAACCTG 180
 QY 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 181 GCTTCGTTTCGACGACGCTGGATGGCCGAGTGCATCGTCTGATGATGTTCCACCATC 240
 QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 Db 241 AACAGAACTACGCCCTCATGGACGATTCACCGGTCACCTATTGACATACAGAAATCGCTGC 300
 QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
 Db 301 GTGAATATGATAGCAACCTCTTTAATCGCCAAATTTGGGAGGGGGAAACAACAGTAGGA 360
 QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 361 TGTGCTACCGTGGGATCATCAGAACCAATGCTTTGCGGGTTGGCAATTCAGAGAAAT 420
 QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 Db 421 TGGCAGACAAAGAAAGGACAGGGGAAAGCCATATGACAAAGCCCAACATGGTCCACCGT 480
 QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 481 TCAAATGTTTCAGGTTTGTGGTGAATTCGCTAAGTATTTTGAAGTTTGAATGAAAAA 540

```

181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
|||
541 GTGAATTTGAGGAGGGATATTATGATGGAGCCAGAGAGGCTGTGAAATGGTGGAT 600
|||
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
|||
601 GAGAAATACCAATTTGTGCTGCTGCACTTGGGCTCGACCTTACTGGAGATTGAGAT 660
|||
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
|||
661 GTCAAACTGTAACACGACCTCCCTCGTAGAGACAGACAGAAAACGTTGGGATACACC 720
|||
241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
|||
721 ATTCATGTTGATGCTGCTATTGGTGGATTCATTTGCTCCATTTATCCAGAACCTG 780
|||
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
|||
781 TGGGATTTCCGACTACTCTGGGTGAAGATATCAATGTCAGTGGACACAAATACGG 840
|||
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 300
|||
841 GTCTATCCCGCGCTTGGTGGTGGTGGTGGGAGAACAGAAATGATCTTCTGAAGAA 900
|||
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
|||
901 ATTTCCATATCAACTATCTTGGGATGATCAACCCACATTTTACCCTCAACTTCT 960
|||
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
|||
961 GGTTCAAACACAGATAATGGTCAACTACTATCAATTAATTCGCCCTTGGTGGG 1020
|||
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
|||
1021 AAGATATAATGGAAAACCTCGACGGAGATGCAAGAAATCTTAGGGAACATCTCG 1080
|||
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
|||
1081 ATGGCGCTTTTCGAGATCATCTCCAGGATATTGGGGCGCTCTCGTCACAATTTG 1140
|||
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTyr 400
|||
1141 AAGCAGCAGCAACAATAGTGTCTTTAGATAGCCGATACAAATAGAGGTTGGATGG 1200
|||
401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
|||
1201 ACAATTCCTGCATACACAATGCCAAAGGAGCTTCAGCACATAGCCCTCTCGTGT 1260
|||
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
|||
1261 ATCAGGAGGACATTCAGCCGGAGCCTCGCCGAGCCCTAGCTAATGACATGAAGAAG 1320
|||
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
|||
1321 TTGGTTGAGCTGGACATACATCCAGTGGACC-----ACCACATGGCCCGCTT 1371
|||
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAsp 479
|||
1372 AAGCCAGTGGAGAAATGGCAATGGTAATATGTC---ATCAAGAAGAGATATTGTAGAG 1425
|||

```

search completed: October 22, 2003, 18:41:35
Job time : 404 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

1 protein - nucleic search, using frame_plus_p2n model

in on: October 22, 2003, 15:16:40 ; Search time 97 Seconds
 (without alignments)
 2284.271 Million cell updates/sec

File: US-10-006-852-2
 Project score: 2615
 Sequence: 1 MWLSHAVSRSDVSVHSTFAS.....DIITGKMKFVADRKKTSGIC 502

Scoring table:
 BLOSUM62
 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

C	13	257	9.8	359	4	US-09-072-596-187	Sequence 187, App
	14	246.5	9.4	1707	4	US-08-939-309-1	Sequence 1, Appli
	15	246.5	9.4	1707	4	US-09-849-180-1	Sequence 1, Appli
	16	246.5	9.4	1707	4	US-09-356-643B-5	Sequence 5, Appli
	17	227	8.7	1707	4	US-08-939-309-3	Sequence 3, Appli
	18	227	8.7	1707	4	US-09-849-180-3	Sequence 3, Appli
	19	227	8.7	1707	4	US-09-356-643B-7	Sequence 7, Appli
	20	225	8.6	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	21	224	8.6	2130	4	US-09-740-369-1	Sequence 1, Appli
	22	177	6.8	3162	4	US-09-356-643B-12	Sequence 12, Appli
	23	160.5	6.1	1467	4	US-08-939-309-9	Sequence 9, Appli
	24	160.5	6.1	1467	4	US-09-849-180-9	Sequence 9, Appli
	25	160.5	6.1	1467	4	US-09-356-643B-9	Sequence 9, Appli
	26	133.5	5.1	1578	4	US-09-328-352-366	Sequence 366, App
	27	121.5	4.6	1830121	4	US-09-557-884-1	Sequence 1, Appli
	28	121.5	4.6	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	29	116	4.4	1944	3	US-09-377-557-19	Sequence 19, Appli
	30	113	4.3	1197	4	US-09-328-352-381	Sequence 381, App
	31	110.5	4.2	129908	4	US-09-585-858-1	Sequence 1, Appli
	32	105	4.0	2190	4	US-09-134-001C-1891	Sequence 1891, Ap
	33	105	4.0	5340	4	US-09-627-122-21	Sequence 21, Appli
	34	105	4.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	35	104	4.0	3312	4	US-09-669-751-259	Sequence 259, App
	36	104	4.0	4032	1	US-08-107-748-3	Sequence 3, Appli
	37	104	4.0	4032	1	US-08-245-809-4	Sequence 3, Appli
	38	104	4.0	4032	5	PCT-US92-01385-3	Sequence 3, Appli
	39	102.5	3.9	2193	3	US-09-398-865A-1	Sequence 1, Appli
	40	102.5	3.9	2193	4	US-09-710-714-1	Sequence 1, Appli
	41	102.5	3.9	2992	1	US-07-718-575-1	Sequence 1, Appli
	42	102.5	3.9	2992	1	US-08-481-206-1	Sequence 1, Appli
	43	102.5	3.9	2992	2	US-08-486-269A-1	Sequence 1, Appli
	44	102	3.9	1734	3	US-08-687-590-60	Sequence 60, Appli
	45	101.5	3.9	3666	4	US-09-107-532A-305	Sequence 305, App

ALIGNMENTS

RESULT 1
 ; Sequence 8, Application US/08522421
 ; Patent No. 5908973
 ; GENERAL INFORMATION:
 ; APPLICANT: Abu-Bakar Umi Kalsom
 ; APPLICANT: Barton, Sarah Louise
 ; APPLICANT: Gallego-Verigas, Pedro Pablo
 ; APPLICANT: Gray, Julie Elizabeth
 ; APPLICANT: Grierson, Donald
 ; APPLICANT: Lowe, Alexandra Louise
 ; APPLICANT: Picton, Steve
 ; APPLICANT: Whotton, Lee Colin
 ; TITLE OF INVENTION: DNA, DNA CONSTRUCTS, CELLS AND PLANTS
 ; TITLE OF INVENTION: DERIVED THEREFROM
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Pillsbury Madison & Sutro, L.L.P.
 ; STREET: 1100 New York Avenue, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3918
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Microsoft Word
 ; CURRENT APPLICATION NUMBER: US/08/522,421
 ; FILING DATE: 11-JAN-1996
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9305868.3
 ; FILING DATE: 22-MAR-1993
 ; PRIOR APPLICATION DATA:

SUMMARIES

result	No.	Score	Query Match	Length	DB ID	Description
	1	2014	77.0	1783	2	US-08-522-421-8
	2	1114	42.6	4403765	3	US-09-103-840A-2
	3	1114	42.6	4411529	3	US-09-103-840A-1
	4	1025	39.2	5565	3	US-09-068-195-21
	5	605	23.1	36941	4	US-08-311-731A-130
	6	325	12.4	1770	4	US-08-939-309-7
	7	325	12.4	1770	4	US-09-849-180-7
	8	325	12.4	1770	4	US-09-356-643B-1
	9	265.5	10.2	1629	4	US-08-939-309-5
	10	265.5	10.2	1629	4	US-08-849-180-5
	11	265.5	10.2	1629	4	US-09-356-643B-3
	12	257	9.8	359	4	US-09-056-556-192

Pred. No. is the number of results predicted by the chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

APPLICATION NUMBER: GB 9305869.1
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9305859.2
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9305865.9
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9305866.7
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9305867.5
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9305860.0
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9305862.6
 FILING DATE: 22-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9314351.9
 FILING DATE: 12-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9320988.0
 FILING DATE: 12-OCT-1993

INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1783 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 ORIGINAL SOURCE:
 ORGANISM: ERTD1
 S-08-522-421-8

Alignment Scores:
 seq. No.: 3.64e-239 Length: 1783
 core: 2014.00 Matches: 387
 percent Similarity: 86.44% Conservative: 53
 est Local Similarity: 76.03% Mismatches: 55
 query Match: 77.02% Indels: 14
 B: 2 Gaps: 5

S-10-006-852-2 (1-502) x US-08-522-421-8 (1-1783)

```

Y 1 MetValLeu---SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19
b 6 ATGGTGTAAACAGCGCTCGATAAGAGATTACAGAGAGCTTGCCTACATTTGCA 65
Y 20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
b 66 TCAAGATATACAGGAACTTACCTAACCTTCAAAATGCTAAATAAATCCATGCCGAAA 125
Y 40 GluAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 59
b 126 GAAGCAGCTTATCAGATGTAAAGACGAGCTATGTTGGATGGTAAACCCAGGTTGAAT 185
Y 60 LeuAlaSerPheValThrTyrMetGluProGluCysAspLysLeuIleMetSerSer 79
b 186 TTASCTTCCITTTGTAGCACATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 245
Y 80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
b 246 ATTAATAAAACTATGTCGACATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 305
Y 100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
b 306 TGTGTAAACATGTAGCACATCTTTCCATGCCCGGCTTGGTGGATGATGAGATGCAGTT 365
Y 120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139

```

```

Db 366 GGAGTTGGTACAGTGGGTTTCATCAGAGCGCAATAAATGCTTGTGTCGCTTCTTCAAACGC 425
Qy 140 LysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThr 159
Db 426 AAATGGCAATCGAAAAGAAAAGCAAGAGGCAAAACCTTTCGATTAAGCCTAATATAGTCACT 485
Qy 160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaIaIaIaIaIaIaIaIaIaIaIaIa 179
Db 486 GGAGCTAATGTGCAGGTTCTGCTGGAAAATAATTCGAAGGTATTTTGGAGTTGAGTTGAAG 545
Qy 180 GluValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetVal 199
Db 546 GAGGTAAACTTAAAGAGGATACATATGTAATGACCCCTCCCAAGCAGTAGAGATAGTG 605
Qy 200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
Db 606 GATGAGAAATACAAATATGTTGCTGCAATCCCTGTTGTTCTACTCTGACTGGGGAGTTGAG 665
Qy 220 AspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThr 239
Db 666 GATGTGAAGCTCCAAAAGGACTCCTTACRAAAAAGAACAAAGGAAACCGGATGGAGACA 725
Qy 240 ProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
Db 726 CCGATTTCATGTCGATGCTGGAGTGGAGGATTTATGCTCCTTTCCTCTGCGCAGATCTT 785
Qy 260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
Db 786 GAAATGGGATTTCCGTTTGCCTCTTGTGAATAAGTATAAATGTCAGCGGTCAACAAGTATGGC 845
Qy 280 LeuValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluGlu 299
Db 846 CTTGTATATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 905
Qy 300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
Db 906 CTCGCTTTTCATATAAACTACCTTGGGTCGATCAGCCTACTTTTACTCTCAACTTCTCT 965
Qy 320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
Db 966 AAAGGTTCCCTTCAATAAATGTCACAGTATTTATCAGTAAATAAGACTGGCTTTGGAGGT 1025
Qy 340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
Db 1026 TATAAAGACGTCATGAAGATTTCTTAAACCCAAAAGTACTAACAGGGAATCACA 1085
Qy 360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
Db 1086 AAAATGGCGCGGTTTCGATATTGCTCTAAAGGATGGGGTGTCTCTGTGTAGCATTCTCT 1145
Qy 380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
Db 1146 CTCAGGGACAGCAGCAAAATATACGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGA 1205
Qy 400 TrpIleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgVal 419
Db 1206 TGGATCGTCCCTGTCATACAAATGCCACCGGATGCTGACACACATGCTGTACTGGGGTT 1285
Qy 420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
Db 1266 GTCATTTAGAGAGGATTTCCAGCCACAGCCTAGCTGAGAGACTTGTTTCTGACATTTGAGAAA 1325
Qy 440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle----SerLeu 457
Db 1326 ATTCGTTCAGAGTTGGACACACAGCCTCCTCGTTTGGCCCAAGCTGTCGCTGTCAC 1385
Qy 458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeu-----MetValThrVal 473
Db 1386 GCTGAGGAGAGTCGTGATGACAGGGTGTATGGGCTTCATCATTTTTCATGATGATGATGTA 1445
Qy 474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTyrLysLysPheValAla 493
Db 1446 -----GAGACTCAGAAAAGACATTTCAAACATTTGGAGAAAATCGCAGG 1490

```

```

y 494 AspArgLysLysThrSerGlyLeuCys 502
b 1491 -----NAGAAACCAGCGGAGTCTGC 1511

```

```

RESULT 2
S-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
S-10-006-852-2

```

```

Alignment Scores:
red. No.: 6.33e-122 Length: 4403765
core: 1114.00 Matches: 219
Percent Similarity: 65.63% Conservative: 77
Local Similarity: 48.56% Mismatches: 143
Query Match: 42.60% Indels: 12
B: 3 Gaps: 5
S-10-006-852-2 (1-502) x US-09-103-840A-2 (1-4403765)

```

```

y 6 AlaValSerGluSerAspValSerVal-----HisSer-----ThrPheAlaSer 20
b 3837651 GCCGTGTCTCCGAGTCAACCCTCCGTCGCGCAATCGATCGCTCCGCGCTACACCGGT 3837592
y 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
b 3837591 CGGATGTCTCCGACCGCGTCCGGCCTCCGGATGCCGATGATGGATCCGGAG 3837532
y 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
b 3837531 CGCGCTATCGCTTATCCACGACGAGCTGATGCTGACGGTAGCTCGCGGCTGAATCTG 3837472
yy 61 AlaSerPheValThrThrMetGluProGluCysAspLysLeuIleMetSerIle 80
bb 3837471 GCCACCTCTGTGACCACTGATGGACCGGACCGGAGCCGAAAGTGTATGGCCGAGCGTTC 3837412
yy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
bb 3837411 GACAAGAATCATGATCGATGATAGGAGGAATACCCCGCGACCGCGCATCGAGGGCGGTGT 3837352
yy 101 ValAsnMetIleAlaHisLeuPheAsnAlaPro-----LeuGluGluAlaGluThr 117
bb 3837351 GTGTCCATGTTCCGCGAGCTTTTTCACGCCGAGGCTGTGGCGCACACCGCCACCGAGC 3837292
yy 118 AlaValGlyValGlyThrValGlySerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
bb 3837291 GCCACCGGGGTCCACCATCGCTCCAGCGAGGGGTGTATGCTGGGTGGGCTGGCCCTG 3837232
yy 138 LysArgLysTrpGlnAsnLysArgLysAla---GluGlyLysProValAspLysProAsn 156
bb 3837231 AAAATGGCGTGGCGGACCGGGTCTTGTAGAGGGGGCGC-----ATGCCCAAT 3837181

```

RESULT 3

```

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A

```

```

157 IleValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluVal 176
Db 3837180 CTGGTGTATGTTCCAAAGTCCAGTGTGGAGAAAGTTCTGCCCGTACTTCCAGCGTC 3837121
Qy 177 GluLeuLysGluValLysLeuSerGluGlyTyrValMetAspPrgGlnGlnAlaVal 196
Db 3837120 GAACCCCGTTACCTGCGCATGGAGGGCGCGCTACGTCATCACCCCGGACAGTGTCTC 3837061
Qy 197 AspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGly 216
Db 3837060 GCCGCGCTCGACGAGAACACCATCGCGTGTGGCGATCTTGGCGCACCCATACCGGT 3837001
Qy 217 GluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGly 236
Db 3837000 GAACCTCGAAACCCATCGCGAGATCTGCGCCGCTGGACAACACTGGCGGCTGGCGGG 3836941
Qy 237 TrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyr 256
Db 3836940 GTGGACGTTCCCGGTACAGCTCGACGCGCCAGTGGGGCTTTGGTCCCTTTTTCAT 3836881
Qy 257 ProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis 276
Db 3836880 CCGGACCTGGTATGGGATTTTCGGCTGCCCGCTGTGTGATCAACAGTCAAGCGGCCAC 3836821
Qy 277 LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTyrArgAsnLysGluAspLeu 296
Db 3836820 AAGTATGGCTGACCTATCCCGGGCTGGGTTTGTGGTGGCGGGCCCGAGCACCTG 3836761
Qy 297 ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
Db 3836760 CCGGAGGATCTGGTTTTCCGGGTCAACTACCTCGCGCGGACATGCCGACCTTCAACCTG 3836701
Qy 317 AsnPheSerLysGlySerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly 336
Db 3836700 AACTTCTCCGCTCCCGTAAACAGCGTGGTGGCCAGTACTACACTTCTCGCGGTGGG 3836641
Qy 337 HisGluGlyTyrArgAsnValMetGluAsnGlyArgGluAsnMetIleValLeuArgGlu 356
Db 3836640 CCGGACCGCTATPACCAAGTGTATGCGCGGCTGTCACACCGCCCGGTGGTGGTGTGAC 3836581
Qy 357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
Db 3836580 CAGCTGCGCGAGTGGATCATGCGAGGTGATCTCGGATGGTTCGGCGATCCCGGTGTGTC 3836521
Qy 377 AlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
Db 3836520 AGCTTCCGCTCCCGCGGACCGCGGGTACACGGAGTTCGACGTCTCCCATGAGCTGCGG 3836461
Qy 397 ArgTyrGlyTrpIleValProAlaThrMetProProAsnAlaGlnHisIleThrVal 416
Db 3836460 ACCTTCGGGTGGCGGTCGCCCGCTACACCATGCGGACCAACCGCCAGCCAGCGGCG 3836401
Qy 417 LeuArgValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAsp 436
Db 3836400 CTGCGGATCGTGGTTCGGGAAAGACTCTCCGCGCAGTTCGGCGGCGCCCTGCACGACGAC 3836341
Qy 437 IleGluLysValMetArgGluLeuAspGluLeu 447
Db 3836340 GCGGTCAACCGCTGGCTGCCCTGGACAAGTCT 3836308

```



```

CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
S-09-103-840A-1

Alignment Scores:
red. No.: 6.35e-122 Length: 4411529
core: 1114.00 Matches: 219
percent Similarity: 65.63% Conservative: 77
est Local Similarity: 48.56% Mismatches: 143
query Match: 42.60% Indels: 12
B: 3 Gaps: 5

S-10-006-852-2 (1-502) x US-09-103-840A-1 (1-4411529)
Y 6 AlaValSerGluSerAspValSerVal-----HisSer-----ThrPheAlaSer 20
b 3851753 GCCGTGTCGCACTACCCGTCGGTCCCGCGCATTCGATCGCTCCGGCTACACCGGT 3851694
Y 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
b 3851693 CGGATGTCACCGCACCGGTCCCGCGCTCCGATGCCGATCGATCGATCCCGGAG 3851634
Y 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
b 3851633 CGGGCCTATCGCTTACCCAGCAGCTGATGCTCGACGGTAGCTCGGGCTCAATCTG 3851574
Y 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
b 3851573 GCCACCTTCGTGACCACTGGATGACCGGAGCCGAAAGCTGATGGCCGAGAGCTTC 3851514
Y 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
b 3851513 GACAAGAACAATGATCGATAAGGACGAATAACCGCGCACCGCGCCATCGAGGCGCGCTGT 3851454
Y 101 ValAsnMetIleAlaHisLeuPheAsnAlaPro-----LeuGluGluAlaGluThr 117
b 3851453 GTGTCCATGTTGCCCGCCAGCTGTTTCACGCCGAGGGTCTCGCCGACACAGCCACCCAGC 3851394
Y 118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
b 3851393 GCCACCGGGTGTCCACATCCGCTCCAGCGAGCGGTGATGTTGGTGGCTGGCCCTG 3851334
Y 138 LysArgLysTrpGlnAsnLysArgLysAla--GluGlyLysProValaLysProAsn 156
b 3851333 AATGGCGTGGCGGCACCGGTGGGTCTTTGAGGGGGCC-----ATGCCCAAT 3851283
Y 157 IleValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluVal 176
b 3851282 CTGGTGTGGTTCGAACCTCCAGTGGTGGGAGAGTCTCCCGCTACTTCGACGTC 3851223
Y 177 GluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaVal 196
b 3851222 GAACCCGTTTACCTCCGATGAGCGGGCGGCTTACGTCATCAACCCCGAGCGAGTGTCTC 3851163
Y 197 AspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGly 216
b 3851162 CCCGCGTTCGACGAGAACCACTCCGGTGGTGGCGATCTTTGGCACCACTATACCCGGT 3851103
Y 217 GluPheGluAspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGly 236
b 3851102 GAATCGAACCCATCCGAGATCTGCCCGCGCTGGCAAACTGGCGGCTGGCGGGGT 3851043
Y 237 TrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyr 256
b 3851042 GTGACGTCCTCCGGTACACGTCAGCGCGCCAGTGGGGCTTTTGTGTGCCGTTTTTGCAT 3850983
Y 257 ProGluLeuGluTrpAspPheArgLeuProValLysSerIleAsnValSerGlyHis 276

```

```

Db 3850982 CCGGACCTGGTATGGGATTTTCGGCTGCCCGCTGGTTCGATCAACAGTCAGGGCCAC 3850923
QY LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeu 296
Db 3850922 AAGTATGGCTGACCTATCCCGGGTTCGGGTTTCGTGTGGCCGGCCGACACCTTCAACCTG 3850863
QY ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
Db 3850862 CCGGAGGATCTGGTTCCTCCGGTCAACTACCTCCGCGCGACATGCGCACCTTCAACCTG 3850803
QY AsnPheSerLysGlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGly 336
Db 3850802 AACCTTCCTCCGCTCCCGGTAAACCAGGTGGTGGCCAGTACTACAACCTTCTCCGCTGGG 3850743
QY HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
Db 3850742 CCGGACGGCTATACCAAGTGTATGACGGCTGTCCGACCCCGCGGTGGCTGGGTGAC 3850683
QY GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
Db 3850682 CAGCTCGCGAGGTGGATCAATTCGAGGTGATCTCGATGGTTCGGGATCCCGGTGTC 3850623
QY AlaPheSerLeuLysAspSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
Db 3850622 ACCTTCGGCTCCCGCGGACCGCGGCTGTCCGACGGAGTTCGACGCTCCCATGAGTCCGG 3850563
QY ArgTyrGlyTyrIleValProAlaTyrThrMetProAsnAlaGlnHisIleThrVal 416
Db 3850562 ACCITTCGGGTGGCAGGTCCCGCTACACCATGCGGACACCGCACCGGAGTGGCGGTG 3850503
QY LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAsp 436
Db 3850502 CTGCGGATCGTGGTTCGGGAAAGGACTCTCCGCGACCTGGCGCGGCGCTGCACGACGAC 3850443
QY IleGluLysValMetArgGluLeuAspGluLeu 447
Db 3850442 GCGGTACCGCGGTGGCTGCCCTGGACAAAGGTC 3850410

RESULT 4
US-09-068-195-21
; Sequence 21, Application US/09068195B
; Patent No. 6140078
; GENERAL INFORMATION:
; APPLICANT: Sanders, Jan W.
; APPLICANT: Ledebor, Adrianus M.
; APPLICANT: Venema, Gerard
; APPLICANT: Kok, Jan
; TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid
; TITLE OF INVENTION: Bacterium, and Its Use in a Lactic Acid Bacterium for
; TITLE OF INVENTION: Production of Desired Protein
; FILE REFERENCE: Sanders-60113/0252227
; CURRENT APPLICATION NUMBER: US/09/068.195B
; EARLIER FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: PCT/EP97/04755
; EARLIER FILING DATE: 1997-08-20
; EARLIER APPLICATION NUMBER: EP 97200744/7
; EARLIER FILING DATE: 1997-03-13
; EARLIER APPLICATION NUMBER: EP 96202444/4
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 5565
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genomic DNA of
; NAME/KEY: mat_peptide
; LOCATION: (3)..(500)
; OTHER INFORMATION: C-terminus of rnhB

```

```

FEATURE:
NAME/KEY: mat_peptide
LOCATION: (1055)..(1922)
OTHER INFORMATION: rggL = gadr
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (2069)..(3577)
OTHER INFORMATION: ofix = gacd
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (3600)..(4997)
OTHER INFORMATION: gadb
FEATURE:
NAME/KEY: mat_peptide
LOCATION: Complement((5078)..(5563))
OTHER INFORMATION: C-terminus of ORF in opposite direction
S-09-068-195-21

Alignment Scores:
red. No.: 8.5e-116 length: 5565
core: 1025.00 Matches: 203
percent Similarity: 63.46% Conservative: 94
Local Similarity: 43.38% Mismatches: 156
Query Match: 39.20% Indels: 15
B: 3 Gaps: 4

S-10-006-852-2 (1-502) x US-09-068-195-21 (1-5565)
18 PheAlaSerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIle 37
3651 TTGGTTCAGAAAGTGAACAGTGGATTTACCTAAATATATAATAGTCTCAACAAATCAAT 3710
38 ProLysGluAlaIaTyrGlnIleAlaHisLeuPheAsnGluMetLeuAspGlyAsnProArg 57
3711 GAGCCTCAGTGGCCATCAGTTAGTTCAAGATCAAACTTATATGAACTTATATGAACTTATG 3770
58 LeuAsnLeuLaserPheValThrThrTyrMetGluProGluCysAspLysLeuIleMet 77
3771 TTAATAATGGCCAACTTCTCAAACTTATATGAACTTATATGAACTTATATGAACTTATG 3830
78 SerSerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln 97
3831 CAGACTTGGAAAAAATGGCAATGCAAAATCAGATATCAAGAACAACTGAAATGAA 3890
98 AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr 117
3891 AACCGTTGGCTCAACATGCTGACCTTTGGAATGCG-----AGTGAATAAGGAAAA 3944
118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetIleAlaGlyLeuAlaPhe 137
3945 ATTTATGGGACTTCGACAAATGGTTCCTCAGAACCTTGTATGCTTGGGGAAATGGCTATG 4004
138 LysArgGlyThrGlnAsnLysArgLysAlaGluLysProValAsp-----LysPro 155
4005 AAGTTTCTGGCGTAGCGAGCGAABAAATAGCCCTAGATATTAATGCGAAAAGCCA 4064
156 AsnIleValThrGlyAlaAsnValGlnValCysTyrGluLysPheAlaArgTyrPheGlu 175
4065 AACTTAGTCAATTCCTCTGGTTATCAAGTTTGGTGGAAAAAATTCGTGTATTATGGGAT 4124
176 ValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAla 195
4125 ATTGAATGAGAGAGTCCCAATGGATGAGAACATATGCAATTCAAATTTGGAAAAAGTG 4184
196 ValAspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsn 215
4185 ATGGATATGTTGATGAATATACGATTCGGATGAGTATGGAAATATGGGAACTTACT 4244
216 GlyGluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThr 235
4245 GGTCTATGATGATCAAAAGCTTGGAAATTTGATTTGAAAGAAATATAATAAACAGACA 4304
236 GlyTyrAspThrProIleHisValAspAlaIaSerGlyGlyPheIleAlaProPheLeu 255

```

```

Db 4305 GACTACAAAAGTTTATTCACGTAGATGCTGTCAGAGGACTTTATGCTCCCTTTGTT 4364
Qy 256 TyrProGluLeuGluTyrPheArgLeuProLeuValLysSerIleAsnValSerGly 275
Db 4365 GAGCCAGAACTTGGTGGGATTTCCGTTTGAATAATGTCATTTCAATCAATACITCAGA 4424
Qy 276 HisLysTyrGlyLeuValTyrAlaGlyIleGlyTyrValIleTyrArgAsnLysGluAsp 295
Db 4425 CATAAATATGTTAGTATATCCTGGTGTAGTGGTCTTGTGGCGTGACAAAAAATAT 4484
Qy 296 LeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThr 315
Db 4485 TTACTCTGAAGAGTAAATTTAAAGTAAGTATCTTTGGAGGAGAAATTCACCAATGCGG 4544
Qy 316 LeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeu 335
Db 4545 ATTAATTTTCTCACAGTCTTCTCAATTAATCGGTCAATTAATTAATTTGACGTTAT 4604
Qy 336 GlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArg 355
Db 4605 GGATTTGATGGATATAAAGCTATTTCATGAGAGAACGCGATAAAAGTAGCCATGATTTAGCA 4664
Qy 356 GluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeu 375
Db 4665 GAAGAAATTTGAAAAAACAGAAATGTTGAGATTATGACGATGGGCGCAATTCACCAAT 4724
Qy 376 ValAlaPheSerLeuLysAspSerSerCysHis-----ThrGluPheGluIleSerAsp 393
Db 4725 GTCCTCAAAATTTAAAGAAAAATTCAAAACCGTGTGGAAATCTTTATGATTTGGCAGAT 4784
Qy 394 MetLeuArgArgTyrGlyTyrIleValProAlaIaTyrThrMetProProAsnAlaGlnHis 413
Db 4785 CGTTTATTAAATGAAGGATGGCAAGTGGCTGTTATCCACTTCTTAAAAAATTTGGAAAAT 4844
Qy 414 IleThrValLeuArgValIleArgGluAspPheSerArgThrLeuAlaGluArgLeu 433
Db 4845 GAATTCATTCACGTTTGTAGTAAATTCGACAGATTTCCGGAAATATGGCAATTAACCTAT 4904
Qy 434 ValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHis 453
Db 4905 GTTCAAGATATCAAGAGCAATTTGATGACATAAATAAGGCTCAATTTCTATTTCATCAG 4964
Qy 453 sLysIleSerLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrVa 473
Db 4965 GAACCTGAAAATA-----AAACAATATGGCTTACTCACTAA 5000
Qy 473 lLysLysSerAspIleAspLys 480
Db 5001 AGATAAAGCGATATATCTAAG 5022

RESULT 5
US-08-311-731A-130/C
; Sequence 130, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATION TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P. C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 130:

SEQUENCE CHARACTERISTICS:
LENGTH: 36941 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORGANISM: MYCOBACTERIUM LEPRAE
S-08-311-731A-130

Alignment Scores:
red. No.: 2.27e-62 Length: 36941
core: 605.00 Matches: 175
Conservative: 68 Conservativity: 55.86%
Mismatch: 130 Mismatch: 40.23%
Indels: 67 Indels: 23.14%
Gaps: 4 Gaps: 4

32 MetProGluAsnSerIleProLysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMet 51
10475 CTGCCGAGCAICGATGGATCTCTGAGCGGGTACCGCTTTTATTCAGACTAGTAAATG 10416
10415 GTAGCGGAGTTCTCGGCTCAACTAGCGGTACTTTCGTCACCACTGGGTGGATCCG 10356
71 LuCysAspLysLeuIleMetSerSerIleAsnLysAsnTyrValAspMetAspGluTyrP 91
10355 AGTAAAGATTTGATGCGCGAGAGCTTCGACAAAACATGTTTAGACAGGACGAGTATC 10296
91 roValThrThrGluLeuGlnAsnArgCysVal-AsnMetIleAlaHisLeuPheAsnAla 110
10295 CGCGCGCGCGCC-ATCGACGGCGCTCGGTATAGCAAGCGGACCGGCTCTTATTCAGC- 10238
111 ProLeuGluGluAlaGluThrAla-----ValGlyValGlyThrValGlySer 126
10237 CCTAGGAGCTCAGCTAGCACAAACCCCTCAATACCTGTGGGTGTTCCATATTCGGCTCC 10178
127 SerGlu--AlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTyr-----141
10177 AGGATGGGTGGTATGCTGGGTTGCCTGGCCATAAAATGGCGTGTGTGCGGTAGCTA 10118
142 -----GlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrG 160
10117 GTTAGCAGCGT-----AGCACCCCAATCTGGGATGG 10085
160 lyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysG 180
10084 CGCAGGGGTCCAGGTGTTGGGAGAAATCTCTGTTACTTCCGATGTCGAGC--CGCT 10027
180 luValLysLeuSerGluGlyTyrTyrValMet-AspProGlnGlnAlaValAspMet-Va 199
10026 ATCTGTGATCGATGAGGAGGATGCTACGTCATCACTCCCGAGCAGGATTAATCGAGCCCGT 9967
199 lAspGluAsnThrIle---CysValAlaAlaAspIleLeuGlySerThrLeuAsnGlyGluPh 218

9966 CGACGAGACACTACTTGGTGTGGTGGCCACT- TTGCTGCCACTTTCACCGCGCAACT 9908
218 eGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTyrAs 238
9907 CGAACCCGTCGGGAGCTCTGGCTGGCTGGACCAAGCTGGCAGCCGACAAATGGCTTGA 9848
238 pThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGl 258
9847 CATTGAGTGTCCGTCGGTGGCCGCGGAGGGGTTCCGG- ATGCCCTTCTTGCACCCCGA 9789
258 uLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnVal-SerGlyHisLysT 278
9788 GTTGAAGTGGGATTTCTGGCTGCCCGGAGTGTGATCGATCAATGTPAAGTGGCTATAAAT 9729
278 YrGlyLeuVal-TyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuPro 297
9728 ATGGCTCACCTACCCCGGTGGGATTTGGCGTATGGCGGCAACGAAATATTGGCC 9669
298 GluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsn 317
9668 GAGGACTAGTCTTTTGGTCAACTACCTGGCGCGGATATGTAGACTTTTCG7ATTTGAAC 9609
318 PheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHis 337
9608 TTTTCCGGCTGGACCAACAGTGTAGTCACTACTACACTTTCTGTGGCTGGGACGT 9549
338 GluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu--- 356
9548 AACGGTACTGAAAGTCAATGCAAGCCCTTGTCCGTGACGGTGGCTGGCTGATTGAGTAG 9489
357 -----GlyLeuGluLysThrGluArg 363
9488 CTCGGCGTGGTACTTGTGAGCTGATTCGGTATCGGACTATCGACATCGACACCCGACGG 9429
364 -----PheAsnIleValSerLysAspGluGlyValProLeuValAlaPhe 378
9428 CCATCCTGGTGTATACCTCCTTGTCTGTAGAGACTGGG-----9387
379 SerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyr 398
9386 -----TACACCGAGTTCGATCTCTCCACGAACTCCGACATTC 9348
399 GlyTrpIleValProAlaTyr-ThrMetProProAsnAlaGlnHisIleThrValLeuAr 418
9347 GGTGGCAGTCCCGCTATAAATGTCGCCACGATGCCACTGACATCGATCGTTCGTCGCG 9288
418 gValValIleArgGluAspPheSer 426
9287 CATCGTGGTCTGGGAGGGCTCTCG 9263

RESULT 6

US-08-939-309-7
Sequence 7, Application US/08939309
Patent No. 6423527
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
APPLICANT: Zhou, Jianhui
TITLE OF INVENTION: SPRINGOSINE-1-PHOSPHATE LYASE
TITLE OF INVENTION: POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
TITLE OF INVENTION: METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/939,309
 FILING DATE: 29-SEP-1997
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: David, Maki J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 200116.402
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1770 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 NAME/KEY: CDS
 LOCATION: 1..1767
 S-08-939-309-7

Alignment Scores:
 red. No.: 6 54e-30 Length: 1770
 core: 325.00 Matches: 122
 percent Similarity: 44.53% Conservative: 98
 best Local Similarity: 24.70% Mismatches: 188
 Indels: 86
 Gaps: 20
 S-10-006-852-2 (1-502) x US-08-939-309-7 (1-1770)

Y	3	LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr	22
b	316	TTGAGGGGTACCGTAGAAAAGGAGTCCAAAGGTCAAAACATCGATCGAAGACGAACTA	375
Y	23	ValArg-----ThrSerLeuProArgPhe-----LysMetProGluAsnSerIleProLys	39
b	376	ATTAGATCGGACTCTCAGTAAATGAAATTCACAGTTGCCATCCCAATGGGATACCTCAG	435
Y	40	GluAlaAlaTyrGlnIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn	59
b	436	GATGATGTTATGAGAGCTAAAT-----AAATTGAC	468
Y	60	LeuAlaSerPheValThrTrpMetGluProGluCys-----	72
b	469	GACTTGATACCATCCCAATGGAAGGAGAAAGGTCTCTGGTCCGTTTACACCGGT	528
Y	73	-----AspLysLeuIleMetSerSerIleAsn-----LysAsnTyr	84
b	529	GGTGATGTTGATCCACTTACAAACATCGCATACGAAAATAATGCGTTGCCAATCAA	588
Y	85	ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle	104
b	589	TTACATCCCGATGCTTCTCCGCGTACGTAATAAATGAAATCCGAAGTGGTTCTATGGTT	648
Y	105	AlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGlyValGlyThrVal	124
b	649	TTAAGAATGTTAATGGCCCT-----TCTGATACAGGTTGTGTACCACAACTTCA	699
Y	125	GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTyrGlnAsnLys	144
b	700	GGTGGTACAGAAATCCTTGCTTTAGCATGCTGAGCCCTAAATGATATGCCCTTCAICAT	759
Y	145	ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln	164
b	760	CGT-----GGAAATCACCAGAACCAATAATGCTCCCGTAACCTGGCACAT	804
Y	165	ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer	184
b	805	GCTGGTTCAGAAAGCTGTTATCTATCTTTGGCATGAGCTACCCACGTTGGAGTAGAT	864
Y	185	GluGlyTyrTyrMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle	204
Db	865	CCAAGCATATCAAGTGGACCTGGGAAAAGTGAATAATTCATCAATAAGAACACAAAT	924
Qy	205	CysValAlaAspIleLeuGlySerThrLeuAsn-----GlyGluPheGluAspVal	221
Db	925	TTA-----CTGGTCGGTTCGGTCCAAACCTTCTCATGTTGCGGATGATATT	975
Qy	222	LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle	241
Db	976	GAAGGATTGGGTAAA--ATAGCACAAAATAATA-----CTTCCITTA	1017
Qy	242	HisValAspAlaIleSerGlyPheIleAlaProPheLeu-----TyrPro	257
Db	1018	CACGTCGACAGTTGCTAGGTTCTTATTTTCAATTTATGAAAAGGCGTGGTTACAAA	1077
Qy	258	GluLeuGlu--TyrAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis	276
Db	1078	AATCTGCCATTACTTACTTTAGAGTCCGGGAGTCCCTCAATATCATGTGACACTCAT	1137
Qy	277	LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeu	296
Db	1138	AAATATGATTTGCACCAAAAAGGCTCCTCAGTTATAATGTATAGAAAACAGCGACTTACGA	1197
Qy	297	ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu	316
Db	1198	ATGCAATCAGTATTACGTAATCCCTCTGGACTGGCGGTTATATGGCTCTCCACATTA	1257
Qy	317	AsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly	336
Db	1258	GCAGGGTCCAGGCTGGTCTATTTCTGAGTTGTTGGCCACTATGTCACATGGGT	1317
Qy	337	HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu	356
Db	1318	GAAAATGGGTAC-----ATTGAGTCGTGCCAAGAA--ATAGTCGGTGCAGCAATG	1365
Qy	357	GlyLeuGluLysThrGluArgPheAsnIle-----ValSerLysAspGluGly	372
Db	1366	AGTFTAAAATAATACATCCAGAAAACATCCAGACCTGAATAATAATGGGCAACCTAGA	1425
Qy	373	ValProLeuValAlaPheSerLeuLysAspSerCysHisThrGluPheGluIleSer	392
Db	1426	TATTCAGTCATTTTCATTTCTCAAAGACCTTGAACATACAC-----GAACATATCT	1476
Qy	393	AspMetLeuArgArgTyrGlyTyrIleValProAlaTyrThrMetProAsnAlaGln	412
Db	1477	GACAGTGTCCAAAGAGGCTGGCATTTCAATGCCCTACAAAAGCCG-----	1524
Qy	413	HisIleThrValLeuArgValIleArgGluAspPheSerArgThrLeuAlaGluArg	432
Db	1525	-----GTTGCACATACACATGGCCITCACGAGA---TTGAGCGCTCAT	1563
Qy	433	LeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIle	452
Db	1564	GTTGTGGATGAGATCGCACATTTTACGT-----	1593
Qy	453	HisLysIleSerLeuGlyGlnLysSerGluSerAsnSer	466
Db	1594	-----ACTACCGTCCAAAGAGTTGAAGGCAATCAAATTC	1629

RESULT 7
 US-09-849-180-7
 ; Sequence 7, Application US/09849180
 ; Patent No. 6495359
 ; GENERAL INFORMATION:
 ; APPLICANT: Zaba, Julie D.
 ; ; Zhou, Jianhui
 ; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATASE LYASE
 ; ; POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 ; ; METHODS OF USE THEREFOR
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Seed Intellectual Property Law Group
 ; STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98055
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/849,180
 FILING DATE: 04-May-2001
 CLASSIFICATION: <UNKNOWN>
 ATTORNEY/AGENT INFORMATION:
 NAME: Pepe, Jeffrey C.
 REGISTRATION NUMBER: 46,985
 REFERENCE/DOCKET NUMBER: 200116.402
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1770 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
 |||
 760 CGT-----GGAATCACCGAACCGAAATAATGCTCCCGTAACCTGCACAT 804
 QY
 165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
 |||
 805 GCTGGTTTGACAAAGCTCTTATTACTTTGGCATGAAGCTACGCCACGTGGAGCTAGAT 864
 Db
 185 GluGlyTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
 |||
 865 CCAACGACATATCAAGTGGACCTGGGAAAGTGAATAATTCATCAATAAGACACAAT 924
 Db
 205 CysValAlaAspIleLeuGlySerThrLeuAsn-----GlyGluPheGluAspVal 221
 |||
 925 TTA-----CTGGTCGGTTCCGCTCAAACTTTCCTCATGGTATTGCCGATGATATT 975
 Db
 222 LysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrProIle 241
 |||
 976 GAAGATTGGTAAA-----ATAGCACAAAATAATAA-----CTTCCITTTA 1017
 Db
 242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu-----TyrPro 257
 |||
 1018 CACGTCGACAGTTGCTAGTTCCTTTATTTGTTTATGAAAAAGCTGGTTACAAA 1077
 Db
 258 GluLeuGlu---TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis 276
 |||
 1078 AATCTGCCATTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1137
 Db
 277 LysTyrGlyLeuValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeu 296
 |||
 1138 AAATATGGATTTCACAAAGGCTCGCTAGTATTATATGTTATGAAACAGGACTTACGA 1197
 Db
 297 ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
 |||
 1198 ATGCATCAGATTACGTAATCCTCTGGACTGGCGGTTATATGGCTCCCTCCTACATTA 1257
 Db
 317 AsnPheSerLysGlySerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly 336
 |||
 1258 GCAGGTCACAGGCTGGTCTATTCGTGTTGTTGGCCACTATGGTCAACATGGGT 1317
 Db
 337 HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
 |||
 1318 GAAATGGGTAC-----ATTGAGTCGTGCCAAGAA---ATAGTCGTCGACCAATG 1365
 Db
 357 GlyLeuGluLysThrGluArgPheAsnIle-----ValSerLysAspGluGly 372
 |||
 1366 AAGTTTAAAAAATACATCCAGGAAAACATTCAGACCTGAAATATAATGGCAACCTAGA 1425
 Db
 373 ValProLeuValAlaPheSerLeuLysAspSerCysHisThrGluPheGluLeuSer 392
 |||
 1426 TATTCACTTTCATTTCTTCAAAAGACCTTGAACATACAC-----GAACTATCT 1476
 Db
 393 AspMetLeuArgArgTyrGlyTyrIleValProAlaTyrThrMetProProAsnAlaGln 412
 |||
 1477 GACAGTTGTCGAAAGAGCTGGCATTTCAATGGCTTACAAAAGCCG----- 1524
 Db
 413 HisIleThrValLeuArgValIleArgGluAspPheSerArgThrLeuAlaGluArg 432
 |||
 1525 -----GTTGCACTACACATGGCTTCAACGAGA---TTGAGCGCTCAT 1563
 Db
 433 LeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIle 452
 |||
 1564 GTTGGATGAGATCTGCCCATTTTACT----- 1593
 Db
 453 HisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSer 466
 |||
 1594 -----ACTACGTCGAAGAGTTGAAAGCGCAATCAAATTC 1629
 Db

RESULT 8
 US-09-356-643B-1
 ; Sequence 1, Application US/09356643B
 ; Patent No. 6589666

3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
 |||
 316 TTGAGGGTACCGPAGAAAAGGAAAGTCCAAAGCTCAAAACATCGATCGAAGACGAACTA 375
 b
 23 ValArg-----ThrSerLeuProArgPhe---LysMetProGluAsnSerIleProLys 39
 |||
 376 ATTAGTCGGACTCTCAGTTAATGAAATTCACAGTTGCCATCCCAATGGGATACCTCAG 435
 Y
 40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
 |||
 436 GATGATGTTATTGAAGAGCTAAAT-----AAATGAAC 468
 bb
 60 LeuAlaSerPheValThrTrpMetGluProGluCys----- 72
 py
 469 GACTGTATACACATACCCRAATGGAAGAAAGGAAAGTCTCGTGGCTGTACCAGGT 528
 b
 73 -----AspLysLeuIleMetSerSerIleAsn-----LysAsnTyr 84
 Y
 529 GGTGATGATTTGATCCACTTACAAAACATCGCATACGAAAAAATATTCGGTTGCCAATCAA 588
 bb
 85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
 |||
 589 TTACATCCCGATGCTTCTCCTCGCTAGTAAAATGGAATCCGAAAGTGGTTCCTATGGTT 648
 Db
 105 AlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGlyThrVal 124
 |||
 649 TTAAGAATGTTAATGAGCCCT-----TCTGATACAGGTGTGTGATCCACAACCTTCA 699
 bb
 125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
 |||
 700 GGTGGTACAGAAATCTTTCCTTTAGCATGTCTGAGCCCTAAAATGATGATGCCCTTCAATCAT 759
 bb

Alignment Scores:
 red. No.: 6.54e-30 Length: 1770
 core: 325.00 Matches: 122
 Percent Similarity: 44.53% Conservative: 98
 est Local Similarity: 24.70% Mismatches: 188
 Query Match: 12.43% Indels: 86
 B: 4 Gaps: 20

GENERAL INFORMATION:

APPLICANT: Saba, Julie D.
TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
POLYNUCLEOTIDES AND MODULATING AGENTS AND
METHODS OF USE THEREFOR

SEQUENCE ID: 1770

ORGANISM: S. cerevisiae
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1770)

S-09-356-643B-1

Alignment Scores:

Table with 4 columns: red. No., Length, Matches, Conserved. Row 1: 6.54e-30, 1770, 122, 98

S-10-006-852-2 (1-502) x US-09-356-643B-1 (1-1770)

Table with 3 columns: Y, b, Y. Contains sequence alignment data for LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22 and other sequences.

Table with 3 columns: QY, Db, QY. Contains sequence alignment data for CysValAlaAspIleLeuGlySerThrLeuAsn and other sequences.

RESULT 9

US-08-939-309-5
Sequence 5, Application US/08939309
Patent No. 6423527
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
APPLICANT: Zhou, Jianhui.

ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/939,309
FILING DATE: 29-SEP-1997
CLASSIFICATION: 800
AUTHOR/AGENT INFORMATION:
NAME: David, Maki J.

REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 200116,402
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 1..1626
US-08-939-309-5

Alignment Scores:
red. No.: 1,34e-22 Length: 1629
core: 265.50 Matches: 118
Percent Similarity: 41.41% Conservative: 99
Mismatch: 22.52% Mismatches: 209
Indels: 98
Gaps: 22

3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
229 CTAACCAAGTAAAGGACGAGCTTGGAAAAGTCTGAGAAATTTGGATCGAAGCCGGAA 288

23 ValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAlaAla 42
289 TACTTCACCAATCCCAAGCCATTCAGTTGGAAAGAACTGAAGTCTGCGCTGTCTGCC 348

43 ---TyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
349 ATCTATGATGATTTGGAGGACCCAGCTTTTGGAAAGGA-----AGAGTATCTGGAGCA 402

62 SerPhe-----Val 64
403 GTCTTCAATGAGAAGACGACAGGACGAGGAGATGATGAGGAGGTTTCGGAAAA 462

65 ThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsnLysAsnTyr 84
463 TTTGCGCTGGACCAACCACCTTTGGCCAAAATTTG----- 495

85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
496 -----TTCCCTGGAGTGAGAAATCATGGAGGCTGAAGTTGTCGCAATGT 540

105 AlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGlyValGlyThrVal 124
541 TGTAAATATGATCAATGGA-----GATTCGGAGACATGT---GGAACATATGCAACT 588

125 GlySerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
589 GGTGATCAATTCATTCCTTTGGCGTGGCTGCTCATCGT-----AATCGT 636

145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
637 CTTTGTAAAGAGGAAAGATACACAGAG-----ATGATTGTCCTCATCGTCCAT 690

165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
691 GCAGCGTTCCTCAAAGCTGCCGAATGTTCCGATCAAAGTTCCGCAAGATCCAGATTGAT 750
185 GluGlyTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
751 CCTGTTACTTTCAAAGTAGACCTTGTCAAATGAAAGCGCCAAATTAACAAGAGAACA--- 807
205 CysValAlaAspIleLeuGlySerThrLeuAsn-----GlyGluPheGluAspVal 221
808 TGTATG-----TTAGTTGGATCTGCTCAAACCTTCCATTTGGAACTCTGATGACATT 861
222 LysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
862 GAAGCTATTGGACAGCTAGGACTGAA-----TATGACATCCCAAGTT 903
242 HisValAspAlaAspSerGlyGlyPheIleAlaProPheLeu---TyrProGluLeuGlu 260
904 CATGTTGATGCTTGTCTGGTGGTTCTCTTATCGCAATAAAGAACTTCTTCAATAACAGTAC 1083
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
964 TATGACTTCCGTTCCCTGGTGTATCTTCGATTTCTGCAGATAGTCACAAATACGGACTC 1023
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
1024 GCTCCAAAGGGGTCAFCAGTTCCTTTATCGCAATAAAGAACTTCTTCAATAACAGTAC 1083
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
1084 TTCGTGTGATGCTGATTTGGCAAGGAGTATCTATGATCGCTACGCTATGGAAGGATCACGC 1143
321 GlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGlyHisIleGluGlyTyr 340
1144 GCTGGGCACACATTCACACTTTGCTGGCGCGCAATGCTTATCACGCTCAGGAAAGGATAC 1203
341 -----ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
1204 AAGGGCAATGCTAGAAAGATTTGCACATCAAGAAAG-----ATTAGAAAT 1251
357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
1252 GGACTTTCAAAACATTAAGGAAATCAAAATPACAAAGGCGCAAGTGAT---GTTTGTATTGTT 1308
377 AlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
1309 AGCTGGACAAACCAATGATGGAGTT-----GAACCTACAGATTCATAAATTCATGAAG 1362
397 ArgTyrGlyTrpIleValProAlaTyrThrMetProAsnAlaGlnHisIleThrVal 416
1363 GAAAAACATTTGGCAACTGAAATGGACTTCAAATTCACGCTGGAGTT---CATATCATGCTC 1419
417 -LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAs 436
1420 ACTATGAATCAATA-----CTCATCTCTGGA 1443
436 pIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSe 456
1444 CTGCTGAAAGCTTTCGCTCCGATTCGAGAGCTG-----CAGTTGAG 1485
456 rLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSe 476
1486 TTTGTCAA---GCCACAAACCATCGGAATCCGACAAAGCAAGTGAAGCAGCC 1536
476 rAspIleAspLysGlnArgAspIleLeuThrGlyTrpLysLysPheValAlaAspArgLys 496
1537 AICTACGGACTTG-----CTCAAAGTATTCCAGACCAATCGCTT 1575
496 sLysThrSer 499
1576 GTTCACGAGT 1585

o 1420 ACTATGAATCAT-----TTCCTGGAGTGAAGATCAATGAGGCTGAAGTTGTCGCAATGTGT 540
 y 436 pIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSe 456
 b 1444 CTCGCTGAAGCTTTCGCGGATTCGAGAGCTG-----CAGTTGAG 1485
 y 456 rLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSe 476
 b 1486 TTTGTCAAAA-----GCCCAAACCATCGGAATCCGACAAAGCAAGTGAAGAGCC 1536
 y 476 rAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLy 496
 b 1537 ATCTACGGACTTG-----CTCAAAGTATTCAGACCCGATCGCTT 1575
 y 496 sLysThrSer 499
 b 1576 GTTCACGAGT 1585
 S-09-356-643B-3
 Sequence 3, Application US/09356643B
 Patent No. 6569666
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
 POLYNUCLEOTIDES AND MODULATING AGENTS AND
 METHODS OF USE THEREFOR
 FILE REFERENCE: 200116.402C1
 CURRENT APPLICATION NUMBER: US/09/356.643B
 CURRENT FILING DATE: 1999-07-19
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 1629
 TYPE: DNA
 ORGANISM: C. elegans
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(1629)
 S-09-356-643B-3
 Alignment Scores:
 red No.: 1 34e-22 Length: 1629
 core: 265.50 Matches: 118
 Percent Similarity: 41.41% Conservative: 99
 Best Local Similarity: 22.52% Mismatches: 209
 Query Match: 10.15% Indels: 98
 Gaps: 4
 S-10-006-852-2 (1-502) x US-09-356-643B-3 (1-1629)
 y 3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
 b 229 CTAACCAAGTAAAGGACGAGCTTGAGAAAAGTCTGAGAAATTTGGATCGAGACCGAA 288
 y 23 ValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAlaAla 42
 b 289 TACTTCACTACAAATCCCAAGCCATTCAGTTGGAAGAACTGAAGTACTTCGCCCTGTGTC 348
 y 43 ---TyrGlnIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
 b 349 ATCTATGATGATTTGGAAGGACCGACTTTTGGGAAGG-----AGAGTATCTGGAGCA 402
 y 62 SerPhe-----Val 64
 b 403 GTCTTCATAGAGAGACGACAGGACGAGGATGTATGAGGAGGTTCGGAAMA 462
 y 65 ThrThrTrpMetGluProGluCysAspLysIleIleMetSerIleAsnLysAsnTyr 84
 b 463 TTGCGCTGGACCAACCCTTTGGCCAAATG----- 495
 y 85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104

Db 496 -----TTCCTGGAGTGAAGATCAATGAGGCTGAAGTTGTCGCAATGTGT 540
 Qy 105 AlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGlyValGlyThrVal 124
 Db 541 TGTAAATATGATGAATGGA-----GATTCGGAGACATGT---GGAACATATGTCAACT 588
 Qy 125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
 Db 589 GGTGGATCCATTTCAATTCCTTTGGCGTGGCTCATCGT-----AATCGT 636
 Qy 145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
 Db 637 CTTTTGAAAAGAGGAAAAGTACACAGAG-----ATGATTTGCCATCATCGTCCAT 690
 Qy 165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
 Db 691 GCAGGTTCTTCAAAGCTGCCGAATGTTTCCGTATCAAAGTTCGCAAGATTCACAGTTGAT 750
 Qy 185 GluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
 Db 751 CTTGTTACTTTCAAAGTAGACCTTGTCAAATGAAAGCCGCAATTAACAAGAGAACA--- 807
 Qy 205 CysValAlaAspIleLeuGlySerThrLeuAsn-----GlyGluPheGluAspVal 221
 Db 808 TGTATG-----TTAGTTGGATCTGCTCAAACITTCATTTGGAACCTGTTGATGACATT 861
 Qy 222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
 Db 862 GAAAGCTATTGGACAGCTAGGACTTGA-----TATGACATCCAGTT 903
 Qy 242 HisValAspAlaAlaSerGlyPheIleAlaProPheLeu---TyrProGluLeuGlu 260
 Db 904 CATGTTGATGCTTGTCTGGTGGTTCCTTCCATTCCTGAAGAAAGACGAGATCGC 963
 Qy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 Db 964 TATGACTTCGCTGTTCTGGTGTATCTTCGATTTCTGCAGATAGTCAACAATACGGACTC 1023
 Qy 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 Db 1024 GCTCCAAAAGGGTCAATCAGTTGTTCTTATCCCAATAGGAACCTTCTTCAATAATCAGTAC 1083
 Qy 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 Db 1084 FTCTGTGATGTGTGGCAAGAGGTATCTATGCATCGGCTACTATGGAAGGATCAGCC 1143
 Qy 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 Db 1144 GCTGGGCACACATTCGACTTTGCTGGCCGCAATGCTTTATCAGCTCAGGAAGGATAC 1203
 Qy 341 -----ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
 Db 1204 AAGGCCAATGTGAAAGATTTGTGACACTACAAGAAAG-----ATTAGAAT 1251
 Qy 357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
 Db 1252 GGACTTTCAAACITPAGGAAATCAAAATCAAGGCCCAAGTGT---GTTTGTATGTT 1308
 Qy 377 AlaPheSerLeuLysAspSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
 Db 1309 AGCTGGACAAACCAATGATGGAGTT-----GAACTCTACAGATCCATAACTTCATGAAG 1362
 Qy 397 ArgTyrGlyTrpIleValProAlaTyrThrMetProAsnAlaGlnHisIleThrVal 416
 Db 1363 GAAAAACATTCGCAACTGAATGGACTTCAATTCACGCTGGAGTT---CATATCATGCTC 1419
 Qy 417 -LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAs 436
 Db 1420 ACTATGAAATCA-----CTCATCTCTGGA 1443
 Qy 436 pIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSe 456
 Db 1444 CTCGCTGAAGCTTTCGTCGCGGATTCGACAGCTG-----CAGTTGAG 1485

y 456 rLeuGlyGlnGluYSerGluSerAsnSerAspAsnLeuMetValThrValIleGlyLeuGlu 476
 b 1486 TTTGTCACAAA-----GCCACAAACCATCGGAATCCGACAAAGCAAGTGAAGCAGCC 1536
 y 476 rAspIleAspLysGlnArgAspIleThrGlyTrpLysLysPheValAlaAspArgLys 496
 b 1537 ATCTACGGACTTG-----CTCAAAGTATTCAGACCCGATCGCTT 1575
 y 496 sLysThrSer 499
 b 1576 GTTCACGAGT 1585

RESULT 12
 S-09-056-556-192/c
 Sequence 192, Application US/09056556
 Patent No. 6350456
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Dillon, Davin C.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
 NUMBER OF SEQUENCES: 241
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/056,556
 FILING DATE: 07-APR-1998
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Maki, David J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 210121.457
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 192:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 359 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 JS-09-056-556-192

Alignment Scores: Length: 359
 Pred. No.: 1.17e-22
 Score: 257.00
 Percent Similarity: 61.32%
 Best Local Similarity: 41.51%
 Query Match: 9.83%
 Indels: 0
 Gaps: 4
 US-10-006-852-2 (1-502) x US-09-056-556-192 (1-359)
 y 280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
 b 359 CTGACCTATCCCGGGCTGGTTGCTGTGGCGGGCCGAGACCTCGCCGAGCAT 300
 y 300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
 b 299 CTGGTTTCCGGGTCAACTACTCTCGGGGGGACATGCGGACCTTCACCTTCACTTCTCC 240
 y 320 LysGlySerSerGlnValIleAlaGlnTyrTrpGlnLeuIleArgLeuGlyHisGluGly 339

Db 239 CGTCCCGGTAACACAGTGGTGGGCCAGTACTACAACCTCCCTGGGCTGGGGCCGACGGC 180
 y 340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
 Db 179 TATACCAAGTGTATCCAGCGCTGTCGCACACCCCGCTGGCTGGGTGACACAGTGGCC 120
 y 360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
 Db 119 GAGGTGGATCAATGCGAGGTGATCTCGGATGTTCCGGCATCCCGGTGCTCAGCCCGGAC 60
 y 380 LeuLysAspSerSerCys 385
 Db 59 ACACGCCACGGCCAGTGC 42

RESULT 13
 US-09-072-596-187/c
 Sequence 187, Application US/09072596
 Patent No. 6458366
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Dillon, Davin C.
 APPLICANT: Campos-Neto, Antonia
 APPLICANT: Houghton, Raymond
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Twardzik, Daniel R.
 APPLICANT: Lodes, Michael J.
 APPLICANT: Hendrickson, Ronald C.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
 NUMBER OF SEQUENCES: 350
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/072,596
 FILING DATE: 05-MAY-1998
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Maki, David J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 210121.417C9
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 187:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 359 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-072-596-187

Alignment Scores: Length: 359
 Pred. No.: 1.17e-22
 Score: 257.00
 Percent Similarity: 61.32%
 Best Local Similarity: 41.51%
 Query Match: 9.83%
 Indels: 0
 Gaps: 4
 US-10-006-852-2 (1-502) x US-09-072-596-187 (1-359)
 y 280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299

b 359 CTGACCTATCCCAGGCGCGGCTGGTTCGTTGGCGGGCCGAGCACCTGCGGAGGAT 300

y 300 LeuIlePheHisIleAenTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319

b 299 CTGGTTCGGGTCACTACTCCGGGGGACATGCCGACCTTCACCCCTGAATTCTCC 240

y 320 LysGlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGlyHisGluGly 339

b 239 CGTCCGGTAACCAAGGTGTGGGCGGACTACAACTTCCTGGCGCTGGGCGGACGGC 180

y 340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359

b 179 TATACCAAGTGTATGCCAGCGCTGTCGCCACCCCGCGGTGGGTGACCAAGTGC 120

y 360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379

b 119 GAGTGGATCATTGCGAGGTATCTGGATGGTTCGGCGATCCGGGTTCAGCCCCGAC 60

y 380 LeuLysAspSerSerCys 385

b 59 ACAGCCACCGCCAGTGC 42

RESULT 14

S-08-939-309-1

Sequence 1, Application US/08939309

Patent No. 6423527

GENERAL INFORMATION:

APPLICANT: Saba, Julie D.

APPLICANT: Zhou, Jianhui

TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE

TITLE OF INVENTION: POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND

TITLE OF INVENTION: METHODS OF USE THEREFOR

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESSES: 10

ADDRESS: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/939,309

FILING DATE: 29-SEP-1997

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: David, Mark J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 200116.402

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1707 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1704

US-08-939-309-1

Alignment Scores:

Pred. No.: 3.27e-20 Length: 1707

Score: 246.50 Matches: 100

Percent Similarity: 40.84% Conservative: 76

Best Local Similarity: 23.20% Mismatches: 210

Query Match: 9.43% Indels: 45

DB: 4 Gaps: 12

US-10-006-852-2 (1-502) x US-08-939-309-1 (1-1707)

Qy 22 TyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAla 41

Db 334 TATGTGAANAAC-----CTGCCCTGTCCAGGTATGGGCACAGCTGAG 375

Qy 42 AlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61

Db 376 GTTCTGGAGACTCAGAGTACAGCTCCATGGATGTTCTCTGCAAGAGGGAAGCC 435

Qy 62 SerPheValThrTrpMetGluProLysAspLysLeuIleMetSerSerIleAsn 81

Db 436 TCAGGAGTGTGTACAATGGGAACCGAAGCTCACGGAGCTCGTGGTTCAGCTTATGGA 495

Qy 82 Lys-----AsnTyrValAspMetAspGluTyrProValThrGluLeuGln 97

Db 496 GAATTACGTGGAGCAATCCACTGCATCCAGATATCTCCCTGGATTCCGGAAAGTTAGAG 555

Qy 98 AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr 117

Db 556 GCAGAAATCGTTAGGATGACTTGTCCCTCCTCAATGGG-----GGACCAGAT 603

Qy 118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137

Db 604 TCCCTGGATGTTGACTTCTGGGGAACGGAAAGCATCTCTGATGGCTGCAAGCTTAC 663

Qy 138 LysArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIle 157

Db 664 -----CGGACTTGGCGTTAGAGAAGGGATCAAAACTCCAGAAAT 705

Qy 158 ValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGlu 177

Db 706 GTGGTCCCGAGAGTCCCATCTGCATTCCGAAAGCAGCTCATATTTGGATGAAG 765

Qy 178 LeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValasp 197

Db 766 ATTTCCCGAGTTGCATGAAAAGAACATGGAGGTG--GATGTGCAGCAATGAAGAGA 822

Qy 198 MetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGlu 217

Db 823 GCCATCTCCAGGAACACAGCTATGCTGTTCTTACCCACAGTTCCTCATGGTGTG 882

Qy 218 PheGluAspValLysLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrp 237

Db 883 ATGGATCTGTCCCAGAGTGGCCCAAGTTAACTGTCAAG-----TAT 924

Qy 238 AspThrProIleHisValAspAlaIleAspGlyPheIleAlaProPheLeu----- 255

Db 925 AAAATCCCACTCCATGTGGATGCTTGTGGGGGGCTCCTCATTTCTTATGGAGAA 984

Qy 256 -----TyrPro---GluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsn 272

Db 985 GCAGGTACCCACTGGAGAAACCAITTTGATTTCCGGTGAAGGTGTGACCAGCATTCA 1044

Qy 273 ValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIleGlyTyrIleTrpArgAsn 292

Db 1045 GCAGATACTATAAGTATGGCTATGCTCTAAAGGTTTCATCATGTGTGATGACTTAAC 1104

Qy 293 LysGluAspLeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnPro 312

Db 1105 GAGAAGTACAGGACGTACCAGTTCCTTTGTGTGGAGACTGGCAAGGGTGTTCACGCA 1164

Qy 313 ThrPheThrLeuAsnPheSerLysGlySerGlnValIleAlaGlnTyrTyrGlnLeu 332

Db 1165 TCTCCAAAGCATAGCTGGCTACGGCTGTGGCACTATTCAGCTGTGGCGGCCCTTG 1224

Qy 333 IleArgLeuGlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIle 352

Db 1225 ATGCACCTCGGTGAGAACCGCTATGTTGAAGCTACAAACAGACTCAAAACTGCTCGC 1284

353 ValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGly 372
 1285 TTCTGAAGTCAGAACTGGAACATCAAAACATCTTCAATTTTCGGT---GATCTCTCAA 1341
 373 ValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSer 392
 1342 TTGTCACTTATCTCTGGGATCAACAGATTTCACATTT-----TACCGACTACTCT 1392
 393 AspMetLeuArgArgTyrGlyTyrIleValProAlaArgThrMetProAsnAlaGln 412
 1393 AATATGATGCTCTGAAGGGTGGAAATTTT---AATCACTGCGAGTTCCCAAGAAGCATT 1449
 413 HisIleThrValLeuArgValIleArgGluAspPheSerArgThrLeuAlaGluArg 432
 1450 CATTCTGCATTACCTAGTAGTACATACTCGG-----AAGCGA 1485
 433 LeuValIleAspIleGluLysValMetArgGlu 443
 1486 GTGGCATCCAGTCTCTAAAGGATATCCGGGAA 1518

RESULT 15

S-09-849-180-1
 Sequence 1, Application US/09849180
 Patent No. 6495359
 GENERAL INFORMATION:

APPLICANT: Saba, Julie D.
 Zhou, Jianhui

TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 METHODS OF USE THEREFOR

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
 ADDRESS: Seed Intellectual Property Law Group
 STREET: 701 Fifth Avenue, Suite 6300
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98055

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/849,180
 FILING DATE: 04-May-2001
 CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Pepe, Jeffrey C.
 REGISTRATION NUMBER: 46,985
 REFERENCE/DOCKET NUMBER: 200116.402

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1707 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

NAME/KEY: CDS
 LOCATION: 1..1704

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 JS-09-849-180-1

Alignment Scores:
 Score: 3.27e-20 Length: 1707
 Percent Similarity: 246.50 Matches: 100
 Local Similarity: 40.84% Conservative: 76
 Query Match: 23.20% Mismatches: 210
 Indels: 45
 Gaps: 4

US-10-006-852-2 (1-502) x US-09-849-180-1 (1-1707)

QY 22 TyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAla 41
 Db 334 TATGTGAAACT-----CTGCCTGCTCAGGTTATGGGCACACTGAG 375
 QY 42 AlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
 Db 376 GTTCTGGAGAGACTCAGGAGTACAGCTCCATGGATGTTCTCTGGCAGAAGAAAGCC 435
 QY 62 SerPheValThrThrTyrMetGluProGluCysAspLysLeuIleMetSerSerIleAsn 81
 Db 436 TCAGGAGCTGTGTCATATGGGAAACCGAAGCTCAGGAGCTGCTGGTCAGGCTTATGGA 495
 QY 82 Lys-----AsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln 97
 Db 496 GAATTCACGTGGAGCAATCCACTGCATCCAGATATCTTCCCTGGATTCCGGGAAGTTAG 555
 QY 98 AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr 117
 Db 556 GCAGAAATCGTTAGTAGTACTTGTTCCTCTTCAATGGG-----GGACCAGAT 603
 QY 118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
 Db 604 TCCTGTGGATGTGACTTCTGGGGAACGGAAACATCTGTATGCTGCTGCAAGCTTAC 663
 QY 138 LysArgLysTyrGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIle 157
 Db 664 -----CGGACTTGGCGTTAGAGAAGGGGATCAAACTCCAGAAAT 705
 QY 158 ValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGlu 177
 Db 706 GTGGTCCCGAGAGTCCCATGCTCATTCGACAAAGCAGCTATTATTTGGGATGAG 765
 QY 178 LeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAsp 197
 Db 766 AITGTCGGAGTTGCACTGAAAAGAACATGGAGGTG---GATGTCAGCAATGAGAGA 822
 QY 198 MetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGlu 217
 Db 823 GCCATCTCCAGGAACACAGCTATGCTGCTGTTCTACCCACACAGTTTCTCATGGTGTG 882
 QY 218 PheGluAspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrp 237
 Db 883 ATGGATCTCTGCCCGAAGTGGCCAAAGTTAACTGTGAGA-----TAT 924
 QY 238 AspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu----- 255
 Db 925 AAATCCCACTCCATGTGGATGCTGCTGGGGGGCTTCCCTCATTTCTTCAAGGAAA 984
 QY 256 -----TyrPro---GluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsn 272
 Db 985 GCAGGTACCCACTGGAGAAACCAATTTGATTTCCGGGTGAAAGGTGTGACCAGCAATTCA 1044
 QY 273 ValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsn 292
 Db 1045 GCAGATACTCATTAAGTATGGCTATGCTCCTAAAGGTTTCACTAGTCGTGATGACTTAA 1104
 QY 293 LysGluAspLeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnPro 312
 Db 1105 GAGAAGTACAGGACTGACAGTCTTTGTTGGTCAGACTGGCAAGTGGTGTCTTACCGCA 1164
 QY 313 ThrPheThrLeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTrpGlnLeu 332
 Db 1165 TCTCCAAGCATAGCTGGCTCACGGCTGGTGGCATATTGGACCTGTTGGCGGCGCTTG 1224
 QY 333 IleArgLeuGlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIle 352
 Db 1225 ATGCACCTCGGTGAGAACCGGCTATGTTGAAGTACTCAAAACAGATCATCAAAACTGCT 1284
 QY 353 ValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGly 372

```

o 1285 TTCCGAAAGTCAGAACTGGAAAAACATCAAAAACATCTTCATTTTCGGT---GATCCTCAA 1341
y 373 ValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSer 392
o 1342 TTGTCAGTTATTGCTCTGGGATCCACGATTTGACATT-----TACCCACTATCT 1392
y 393 AspMetLeuArgArgTyrGlyTyrIleValProAlaTyrThrMetProProAsnAlaGln 412
b 1393 AATATGATGCTGCTAAGGGGTGGAAATTT--AACTACCTGCAGTTCCCAAGAGCATT 1449
y 413 HisIleThrValLeuArgValIleAaGluAspPheSerArgThrLeuAlaGluArg 432
b 1450 CATTCTGCATTACGTTAGTACATCTCGG-----AAGCGA 1485
y 433 LeuValIleAspIleGluLysValMetArgGlu 443
b 1486 GTGGCGATCCAGTTCCTAAAGGATATCCGGAA 1518

```

earch completed: October 22, 2003, 18:12:40
ob time : 2696 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

MM protein - nucleic search, using frame_plus_p2n model

run on: October 22, 2003, 15:08:39 ; Search time 358 Seconds
(without alignments)
3785.246 Million cell updates/sec

title: US-10-006-852-2
efrect score: 2815
equence: 1 MVLSHAVSESDVSHSTFAS.....DIITGKKFVDRKKTSGIC 502

coring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

searched: 2552756 seqs, 1349719017 residues
otal number of hits satisfying chosen parameters: 5105512

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+p2n.model -DEV=xip
Q=/cgn2_1/USPTO.spool/p/US10006852/runat_21102003_164103_2443/app_query.fasta_1.647
DB=N_Geneseq_19Jun03 -CFAST=fastap -SUFFIX=eng -MINMATCH=0 1 -LOOPEL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human4.0.cdi
LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
MODE=LOCAL -OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
USBR=US10006852@cgn_1_490@runat_21102003_164103_2443 -NCPU=6 -ICPU=3
NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match %, Length, DB, ID, Description. Contains 45 entries of sequence matches and their descriptions.

ALIGNMENTS

RESULT 1
AAL43410
ID AAL43410 standard; cDNA; 1509 BP.
XX AAL43410;
XX
XX
XX 25-SEP-2002 (first entry)
DT
DT
DE A thaliana GAD1 coding sequence.
XX
XX GAD; plant GABA production regulation; glutamic acid decarboxylase;
XX plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
OS Arabidopsis thaliana.
XX
XX Key Location/Qualifiers
XX CDS 1..1509

/*tag= a
/product= "GAD1"

WO200238736-A2.

16-MAY-2002.

07-NOV-2001; 2001WO-US47447.

07-NOV-2000; 2000US-246367P.

(EMER-) EMERALD BIOAGRICULTURE CORP.

Kimnersley AM, Turano FJ;

WPI; 2002-490073/52.

P-PSDB; AAO15132.

Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome

Claim 18; Page 53; 63pp; English.

The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the A. thaliana GAD1 coding sequence.

Sequence 1509 BP; 433 A; 321 C; 385 G; 370 T; 0 other;

Alignment Scores:

Record: No.: 2,38-262 Length: 1509
Score: 2615.00 Matches: 502
Percent Similarity: 100.00% Conservative: 0
Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 24

US-10-006-852-2 (1-502) x AAL43410 (1-1509)

1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
1 ATGGTGTCTCCACGCGGTATCGGAGTGGACGCTCCCGTCCACTCCACTCCGATCA 60
21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
61 CGTTACCTCCGACTTCCATCTCCCTAGGTTCAAGATGCGGAAACTCGATTCTAAGGAA 120
41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetIleuAspGlyAsnProArgLeuAsnLeu 60
121 GCGCGGTATCAGATCATCAACGAGCGACTGATGTTACCGGAAATCCACGGTGAACCTA 180
61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
181 GCCTCCCTTGTGACGACATGGATGAGCTGATGTGATTAACCTCATCTCTCCCTCAATC 240
81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
241 AACCAAGAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
301 GTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140

361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTCGCTTCAAGCGTAAA 420
141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnLysValThrGly 160
421 TGGCAGACAACGCGCAAGCTGAAGCAAAACCCGTCGATAAACCACCACTTCCACCGA 480
161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
481 GCCAATGTTCAAGTGTGGGAGAAATTCGCTAGTACTTTGAGGTTGAACTTAAGGAA 540
181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
541 GTGAATTTGAGTGAAGGATACATGATGATGATGATGATGATGATGATGATGATGATG 600
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
601 GAGAACACCATTTGTTGGCGGACATCTTGTTCCTTCCACTTTAATGGAGAAATTCGAAG 660
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAAACAAGAAACCCGGATGGGATACA 720
241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
721 ATCCACGTGGATGCGGCAAGTGGAGATTCATTTGCCCGTTTTGTTATCCGGAATTTG 780
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
781 TGGGACTTTAGACTTCCCTTGGTGAAGATGATCAATGTGAGTGTGTCACAAATATGACT 840
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeu 300
841 GTGTACCGAGGATGTTGGTGTGATCTGGAGAAACAAGAGGATTTGCTGAGAACTC 900
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
901 ATCTTCCATATCAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
961 GGTTCAAGTCAAGTCTCATCTCAATCACTACCAACTATCCGATGGCCACCGAGGTTAC 1020
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuLys 360
1021 AGAAATGTGATCGAATTCAGAGAGAAATATGATGATGATGATGATGATGATGATGATG 1080
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
1081 ACAGAAAGGTTCAACATCGTCTCAAAAGGACGAGGAGTCCCACTTGTCCGTTTCCCT 1140
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
1141 AAAGATGACGCTGTCCACTGAGTTCGAAAATCTCCGACATGCTTCGACAGTATGGATGG 1200
401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
1201 ATAGTCCCGCCCTTACCAATGCTCCAAATGCTCAACACACATCACCTGTTCTTCGTGG 1260
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGlyLysVal 440
1261 ATCAGAGAAATTTCTCGAAGACACTCCCTGAGAGACTTGTGATCGATATAGAGAAAG 1320
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1321 ATCGCTGAGCTCGATGAGCTTCCCTCGAGAGTATTCACAAATATCACTTGGCAAGAG 1380
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1381 AAGAGTGAATTCACAGCGGATAACTTTGATGTTGATGTTGATGTTGATGTTGATGTTG 1440
481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
1441 CAGAGAGATATCATCTGCTGGAAGAAGTGTTCGCCGACAGGAAAGACGAGTGGT 1500

```

21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
  |||||
61 CGTTACGTCCTTCACTCCAGTAGTTCAGGAACTCGAATTCCTAAGGAA 120
  |||||
41 AlaAlaTyrGlnIleAsnAspGluLeuMetLeuAspGlyValAsnProArgLeuAsnLeu 60
  |||||
121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTACCGGGAATCCACGGTGAACATA 180
  |||||
61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
  |||||
181 GCCTCTTGTGACGACATGGATGGAGCTGAGTGTGATAAACTCATCATGTCTCCATC 240
  |||||
81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
  |||||
241 AACAAAGAACTATGTTGACATGCAGAGTACCCCGCTACCCCGAACCTTCAGAACCGATGT 300
  |||||
101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluAlaGluThrAlaValGly 120
  |||||
301 GTGAACATGATGGCACATCTATTCAATGCACCGTTAGAGAGGGCGGAGCCGCCCTCGGA 360
  |||||
121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
  |||||
361 GTAGGAACCGTTGGATCATCGAGGCCATAATGTTGGCCGTGGTTCGAAGCGTAAA 420
  |||||
141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
  |||||
421 TGGCAGAAACCGCGCAAAGCTGAAGGCAACCCGTCGATAAACCCCAACTGTACCACCGA 480
  |||||
161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
  |||||
481 GCCAATGTTCAAGTGTGTTGGAGAAATTCGCTAGGTACTTGGAGTTGAATTAAGNA 540
  |||||
181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
  |||||
541 GTGAAATGAGTGAAGGATATGTGTGTGGCCCTCAACAAGCTGTTGATGTGTTGAT 600
  |||||
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
  |||||
601 GAGAACACCATTTGTTGGCCCAATCTTGTTGGTCCACTTAAATGGAGAAATTCGAAAGT 660
  |||||
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrPro 240
  |||||
661 GTTAAACTCTTGAAAGCATCTCTTGGTCCGAAAGAACAAAGAACCCGATGGATACACCA 720
  |||||
241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
  |||||
721 ATCCACGTCGATGCGCAAGTGGAGGATTGATGCACCGTTTTTTGTATCCGGAATTGGAA 780
  |||||
261 TrpAspPheArgLeuProIleuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
  |||||
781 TGGGACATTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCCACCAAGTATGGACTT 840
  |||||
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuLeu 300
  |||||
841 GTGTCAGCGAGGATGTTGGTGTGATCTGGGAAACAAAGAGGATTTCCCTCAGGAACATC 900
  |||||
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
  |||||
901 ATCTTCCATATCAATATCTTGGTGTCCGACCAACCCACTTACTCTCAATTTCTCCAAA 960
  |||||
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
  |||||
961 GGTTCAGTCAAGTCAATGCTCAATPACACCACCTTATCCGATGGGCGCACGAGGTTAC 1020
  |||||
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGluLeuLysLys 360
  |||||
1021 AGAANTGTGATGGAGAATTCAGAGAGHAATGATCGTCCCTCCAGGAGAGGACTTGGAGAG 1080
  |||||
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
  |||||
1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGAGTCCCACTTGCCCTTTCTCTTCC 1140
  |||||
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
  |||||

```

```

Y 501 IleCys 502
b 1501 ATCTGC 1506
ESULT 2
BZL14005
D ABZ14005 standard; DNA; 1509 BP.
X ABZ14005;
X C
X X 21-JAN-2003 (first entry)
E Arabidopsis thaliana stress regulated gene SEQ ID NO 1810.
W Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
X S Arabidopsis thaliana.
X X WO200216655-A2.
X X 28-FEB-2002.
X X 24-AUG-2001; 2001WO-US26685.
X X 24-AUG-2000; 2000US-227866P.
R R 26-JAN-2001; 2001US-264647P.
R R 22-JUN-2001; 2001US-300111P.
X X (Scri ) SCRIPPS RES INST.
X X (SYGN ) SYNGENTA PARTICIPATIONS AG.
X Harper JF, Kreps J, Wang X, Zhu T;
WPI; 2002-304127/34.
X Identifying a stress condition to which a plant cell has been exposed
X and producing plants with increased tolerance to these abiotic stresses
X -
X Claim 144; SEQ ID NO 1810; 577pp + Sequence Listing; English.
X The invention relates to identifying a stress condition to which a plant
X cell has been exposed, comprising:
X (a) contacting nucleic acid representative of expressed polynucleotides
X in the plant cell with an array or probes representative of the plant
X cell genome; and
X (b) detecting a profile of expressed polynucleotides in the plant cell
X characteristic of a stress response. The method is useful in the
X production of transgenic plants, cells and seeds and in producing plants
X with increased tolerance to abiotic stress. The present sequence is that
X of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
X in methods of the invention.
X C Note: The sequence data for this patent is not represented in the printed
X specification but is based on sequence information supplied to Derwent by
X the European Patent Office.
X Sequence 1509 BP; 432 A; 322 C; 385 G; 370 T; 0 other;
Alignment Scores:
Score: 1.57e-261 Length: 1509
Match: 2607.00 Matches: 501
Conservative: 0
Mismatch: 99.80% Mismatches: 1
Indel: 99.69% Indels: 0
Gap: 24 Gaps: 0
US-10-006-852-2 (1-502) x ABZ14005 (1-1509)
1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
1 ATGGTGCTCTCCCAACCGCGATCGAGTGGACGTCTCCGTCCTCCGACATCCGATCA 60

```


1141 AAAGATAGCAGCTGTACACTAGTTCGAAATCCGACATGCTCCGAGTATGGATGG 1200
 401 IleValProAlaIyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1201 ATAGTCCCGCCCTACACAATGCCCTCCAAATGCCAAACACACATCACTGTCTCTCGTGGT 1260
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1261 ATCAGAGAAGATTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisIysIleSerLeuGlnGlu 460
 1321 ATGGTGAGCTCGATGAGCTTCCTTCGAGAGTGAATCCAAAATATCACTTGGACAAAG 1380
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1381 AAGAGTGAATCTAACCGGATAACTTGTGATGGTCAAGGTGAAGAGCGATATCGACAG 1440
 481 GlnArgAspIleIleThrGlyTrpLysPheValAlaAspArgLysLysThrSerGly 500
 1441 CAGAGAGATATCATCACTGGCTGGAGAGATTTGTCGCCCGACAGGAGAGACGAGTGT 1500
 501 IleCys 502
 1501 ATCTGC 1506

RESULT 3
 AAL43415
 D D AAL43415 standard; cDNA; 1705 BP.
 X X AAL43415;
 C X
 T 25-SEP-2002 (first entry)
 X X Tobacco GAD1 coding sequence.
 X X GAD; plant GABA production regulation; glutamic acid decarboxylase;
 X X plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
 X X Nicotiana tabacum.
 H Key Location/Qualifiers
 I CDS 71..1561
 T /*tag= a
 T /product= "GAD1"
 X X MO200238736-A2.
 X D 16-MAY-2002.
 X X 07-NOV-2001; 2001MO-US47447.
 X X 07-NOV-2000; 2000US-246367P.
 X X (EMER-) EMERALD BIOAGRICULTURE CORP.
 X X Kinnersley AM, Turano FJ;
 X X WPI; 2002-490073/52.
 X X P-PSDB; AAOJ5137.

Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 56-57; 63pp; English.

The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a

CC polynucleotide encoding a functional plant glutamic acid decarboxylase
 CC (GAD), to provide a transformed plant that expresses the GAD coding
 CC sequence in response to a signal. Plants of this type have an enhanced
 CC ability to tolerate environmental or other stresses. The present sequence
 CC is the tobacco GAD1 coding sequence.
 XX
 SQ Sequence 1705 BP; 523 A; 301 C; 385 G; 496 T; 0 other;

Alignment Scores:
 Pred. No.: 3,19e-223 Length: 1705
 Score: 2240.00 Matches: 428
 Percent Similarity: 92.05% Conservative: 35
 Best Local Similarity: 85.03% Mismatches: 32
 Query Match: 85.66% Indels: 8
 Gaps: 24
 Gaps: 3

US-10-006-852-2 (1-502) x AAL43415 (1-1705)
 QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 71 ATGGTCTCTCCAAAGACACGCGTCCGAAAAGTACGCTCTCCATCCACTCCACTTCGCTTCC 130
 QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 131 CGRTATGTTCTGCTTCTTCCGAGGTTTAAAGTCCAGAGAAATTCGATACCAAGGAA 190
 QY 41 AlaAlaTyrGlnIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 Db 191 GCAGCATATCAAAATCATAAATGATGAGCTTATGTTAGATGGAAAATCCAAAGTAAATTTA 250
 QY 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 251 GCATCTTTTGTGACAAACATGGATGGAAACAGAGTGTAAACAACTGATGCGATTCCATT 310
 QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 Db 311 AACAAAGATTACGTTGACATGGATGAAATACCTGTAAACCACTGAACTTCAGAAATCGATGT 370
 QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluAlaGluThrAlaValGly 120
 Db 371 GFAAACATGATAGCTCAATTTGTTAAACGACACACTGGAGATGGAGAGACTGCACGTGGA 430
 QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 431 GTTGAACCTGTTGGATCTCTGAGCTATTAATGCTTGTGGATAGCTTTCAAGAGAAA 490
 QY 141 TrpGlnAsnLysArgLysAlaGluLysProValAspLysProAsnIleValThrGly 160
 Db 491 TGGCAAAATATAATGAAAGCCCAAGCAAGCCCTGTGACAGCCCAATATTGCTACTGGT 550
 QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 551 GCCAATGCCAGTGTGTTGGAGAAAATTTGCAAGGTATTTTGAAGTGGAGTAAAGGAA 610
 QY 181 ValLysLeuSerGluGlyTyrTrpValMetAspProGlnGlnAlaValAspMetValAsp 200
 Db 611 GFAAAGTTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
 QY 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 Db 671 GAGAACACAAATTTGTGAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAGAT 730
 QY 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 Db 731 GPTAAGCGCTTGATGACTCTTGATTTGAGAAGAAACAAAAGAAACCGGGTGGGACACTCCA 790
 QY 241 IleHisValAspAlaAlaSerGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 Db 791 ATTCATGTGGATGACAGCAAGTGGTGGATTTATTTGCCCAATCTCTTATCCAGAGCTTGA 850
 QY 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 Db 851 TGGGACTTTAGTTCCCAATTTGGTGAAGAGTATAACCGTGGTGGTTCACAAAATATGGTCTT 910

Y 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 b 911 GTTATGCTGGTATTGGTGGCCATTTGGAGGAATAAGGAAGACTTACTCGCAACTT 970
 Y 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 b 971 ATCTTCCACATTAATATCTGTGGTGTGATCAACCTACTTCTCACTCACTCTTCAA 1030
 Y 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 b 1031 GGTTCAGCCCAAGTAATGCTCAATATACCACTTATTCGCTTGGGTTTGGGGTTAC 1090
 Y 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 b 1091 AAGAAATGTTAGGAAATGCTCAAGAAATGCAAGGGTACTAAGAGAAAGACTTGA AAAA 1150
 Y 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 b 1151 AGTGGAAAGATTCACATPAATCCAAAGAAATGGAGTTCCATTAGTAGCTTCTCTT 1210
 Y 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTyr 400
 b 1211 AAAGACAACAGTCAACACAATGAGTTCGAAATTTTCGAAACTCTTAGAAGATTTGGATGG 1270
 Y 401 IleValProIleTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 b 1271 ATTATTCCTGCATATCTATSCCAAAATGCTCAACATGTCACAGTCTCAGAGTTGTC 1330
 Y 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 b 1331 ATTAGAAGAATTTCTCCCGTACTCGCCGAGGACTGTAATAGACATTTGAARAAGTC 1390
 Y 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGluGlnGlu 460
 b 1391 CTCCACGAGCTAGACACACTTCGCGGAGGGTCAACGCTAAGCTAGCCGFG----- 1441
 Y 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 b 1442 --GCCAGGCGAATGCGCGCGTGCAT-----AAGAAACAGATAGAGAAGTG 1489
 Y 481 GlnArgAspIleIleThrGlyTyrLysPheValAlaAsp--ArgLysLysThrSer 499
 b 1490 CAGCTTGAGATTACTACTGATGGAGAAATTTGTTCTGATAGAAAGAAAGACTAAC 1549
 Y 500 GlyIleCys 502
 b 1550 GGAGTTGT 1558
 RESULT 4
 AAL43417
 standard; DNA; 1784 BP.
 AAL43417;
 25-SEP-2002 (first entry)
 Petunia GAD coding sequence.
 GAD; plant GABA production regulation; glutamic acid decarboxylase;
 plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
 Petunia sp.
 WO200238736-A2.
 16-MAY-2002.
 07-NOV-2001; 2001MO-US47447.
 07-NOV-2000; 2000US-246367P.
 (EMER-) EMERALD BIOAGRICULTURE CORP.

XX Kimmersley AM, Turano FJ;
 WPI; 2002-490073/52.
 P-PSDB; AAO15139.
 Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 58; 63pp; English.
 The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the petunia GAD coding sequence.
 SQ Sequence 1784 BP; 550 A; 298 C; 400 G; 536 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,248-221 Length: 1784
 Score: 2225.00 Matches: 427
 Percent Similarity: 92.45% Conservative: 38
 Best Local Similarity: 84.89% Mismatches: 34
 Query Match: 85.03% Indels: 5
 Gaps: 2
 US-10-006-852-2 (1-502) x AAL43417 (1-1784)
 QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 72 AIGGTTCTATCAAAACAGTGTCCGAGCGGATGTCCATTCCTCCACGTTTCTCT 131
 QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 132 CGATATGTTGCAACTTCTCTCCAGGTTTAAATGCCAGATAATTCGATACCAAAAGAA 191
 QY 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 Db 192 GCAGCATATCAGATCAATAATGATGAATGATGTAGATGGAAACCCCAAGGCTGAAC 251
 QY 61 AlaSerPheValThrThrTropMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 252 GCTTCTTTTGTACAACTGGAATGGAACCGAGTGTGATGATGATGATGATGATGAT 311
 QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
 Db 312 AACAAAGAACTATGTTGATATGGATGAAATATCTGTACCACCTGAGCTTCAGAAATCGATGT 371
 QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
 Db 372 GTAACATGATAGCTCAITTTGTTAATGCCACCTTGAAGATGGAGAACTGCAGTTGGA 431
 QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 432 GTTGGAACTGTTGGATTCCTCTGAAACCAATATGCTTGTGATAGCTTTCAGAGAAA 491
 QY 141 TrpGlnAsnLysArgLysAlaGlyLysProValAspLysProAsnIleValThrGly 160
 Db 492 TGGCAGAACAAAATGAAAGCCCAAGCAACCCCTGTGACAAAGCCCAACATTTACTGGT 551
 QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 552 GCAAAATGCTCAGGTGTGGGAGAAAATTTGCAAGGTATTTTGAAGTGGAGTAAAGAAA 611
 QY 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValasp 200

```

612 GTAAGCTTAGTGAAGGATAGTACTGATGATGAGCCCTGAGAAAGCTGTGGAGATGGTGGAT 671
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
|||
672 GAAACACCACTTGTGTAGCTGCTATCTTAGTTCACCCCTCAATGGAGAAATTTGAAGAC 731
|||
221 ValIysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
|||
732 GTTAAAGCGCTTGAATGATCTTGTGGTCGAGAAGAACAAAGAAAACCGGGTGGACACTCCA 791
|||
241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTrpGluLeuGlu 260
|||
792 ATTCATGTGATGCAGCAAGTGGTGGATTTATGACCCGTTCAATACCACGAGCTTGAG 851
|||
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
|||
852 TGGGACTTATAGATGCCATTAAGTGAAGAGCAATTAAGTAAAGTGTCAAAATATGGTCTT 911
|||
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
|||
912 GTCATGCTGGTATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 970
|||
301 IlePheHisIleAsnTyrLeuGlyValAspGlnProThrPheThrLeuAsnPheSerLys 320
|||
971 ATCTCCACATTAATATCTTGGTGTGATCAACCTACTTCTCAACTTTCTAAA 1030
|||
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
|||
1031 GGTTCAGCCAAAGTAAATGCTCAATATACCACTTATTCGCTGGTGGTGGTGGTGGTGGT 1090
|||
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
|||
1091 AAGAAUUGTATGAGAAATGTCRAAANAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 1150
|||
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
|||
1151 ACAGGAAGATTAACATAAATCCAPAGAAATGGAGTACTTGTAGTACTTCTCTCTT 1210
|||
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
|||
1211 AAAGCAACAGCCAAACACAGCTCCAGATTTCTGAAACTTTAAGCAGATTTGGTTGG 1270
|||
401 IleValProIaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
|||
1271 ATGTTCTCGATATATATGCCAACCAACGCAACACACATTAACAGTTCTCAGAGTTGG 1330
|||
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
|||
1331 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAGCAGCTGTTAGAGACATCGAAAAGTC 1390
|||
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
|||
1391 CTTCAATGAACCTTGACACACTCCCTGCACGCTCAATGCTAAGCTCGTGTGGCCGAGGAG 1450
|||
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
|||
1451 CAGCGCGCTCGAATGCGCAGGAGTGCAT-----AAGAAAACAGATACCGAAGTG 1501
|||
481 GlnArgAspIleIleThrGlyTrpLysPheValAlaAsp---ArgLysLysThrSer 499
|||
1502 CAGTGGAGATGATAACTGCATGGAAGAAGTTGTTGAAGAAAAGAAAGAAAGACTAAT 1561
|||
500 GlyIleCys 502
|||
1562 CGAGTTTGT 1570
|||
RESULT 5
AAL43416
CD AAL43416 standard; cdNA; 1771 BP.
XX
AAL43416;
XX
25-SEP-2002 (first entry)
XX
DE Tobacco GAD2 coding sequence.
XX
KW GAD; plant GABA production regulation; Glutamic acid decarboxylase;
XX
KW plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
XX
OS Nicotiana tabacum.
XX
FH Key Location/Qualifiers
FT CDS 67..1557
FT /tag= a
FT /product= "GAD2"
FN MO200238736-A2.
PD 16-MAY-2002.
XX
PF 07-NOV-2001; 2001WO-US47447.
XX
PR 07-NOV-2000; 2000US-246367P.
XX
(EMBR-) EMERALD BIOAGRICULTURE CORP.
XX
PI Kinnersley AM, Turano FJ;
XX
DR MPI; 2002-490073/52.
XX
DR P-PSDB; AAO15138.
XX
PT Making transformed plants that selectively increase gamma-aminobutyric
PT acid production, by incorporating a DNA construct with a polynucleotide
PT encoding a plant glutamic acid decarboxylase enzyme into plant's genome
XX
PS Claim 18; Page 57; 63pp; English.
XX
CC The present invention relates to a method of producing a transformed
CC plant that selectively increases production of gamma-aminobutyric acid
CC (GABA) in response to a signal, by incorporating into the plant's genome
CC a DNA construct with a non-constitutive promoter operably linked to a
CC polynucleotide encoding a functional plant glutamic acid decarboxylase
CC (GAD), to provide a transformed plant that expresses the GAD coding
CC sequence in response to a signal. Plants of this type have an enhanced
CC ability to tolerate environmental or other stresses. The present sequence
CC is the tobacco GAD2 coding sequence.
XX
SQ Sequence 1771 BP; 548 A; 295 C; 393 G; 535 T; 0 other;
XX
Alignment Scores:
Pred. No.: 6,28e-219 Length: 1771
Score: 2199.00 Matches: 423
Percent Similarity: 90.85% Conservative: 34
Best Local Similarity: 84.10% Mismatches: 38
Query Match: 84.09% Indels: 8
DB: 24 Gaps: 3
US-10-006-852-2 (1-502) x AAL43416 (1-1771)
QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 67 ATGGTTCTGTCACAGACAGCGTCGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 126
QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 127 CGATATGTTGCAACTTCTCTCCAGGTTTAAATGCGCAGAGAATTCATACCAAGAA 186
QY 41 AlaAlaTyrGlnIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 187 GCAGCATATFCAGATATATAATGATGAGCTTATGTTAGATGAAATCCCAAGGCTAAATTA 246
QY 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
Db 247 GCATCTTCGTTACCAACATGGATGGAGCCAGCAATGTAATACGTTAATGATGATCCATT 306

```

81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 307 AACAAAGAACTACGTTGACATGGATGAATACCCCTGTACCACTGAGCTTCAGAAATCGATGT 366
 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 367 GTAATAATGATAGCTCAATTTGTTTATGCACCACTGGAGATGGAGACTGCGATTTGA 426
 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 427 GTTGGAACTGTTGGATCCTCTGAAGCTTATATGCTTCTGATTAAGCTTTAAGAGAAA 486
 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 487 TGGCAAAATAAATAAGAAAGCCCAAGCCCTTTGATAAGCCCAATAATTTGCACCGGT 546
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 547 GCTAAATGTCAGGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAGAA 606
 181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 607 GTAAATTTGAGTGGATGATACTATGATGATGACCCCTGAGAAAGCTGTGGAATTTGGAT 666
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 667 GAGAAATACCAATTTGTGTGCTATCTTAGTTCACACTCACTCAATGGTGAATTTGAAGAT 726
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 727 GTTAAAGGTTTGAATGATCCCTTTGATGAGAGAAACAAGAAACCCGGTGGACATCCA 786
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaPheLeuTyrProGluLeuGlu 260
 787 ATTCAATGATGAGCAAGTGTGGATTTATTCACCACTTCTTATCCAGAGCTTGA 846
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 847 TGGACTTTAGATTGCCAATTTGGAGAAAGTATTAATGATGAGTGCACAAATATGTC 906
 281 ValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeu 300
 907 GTCATGCTGATTTGTTGGCCATTTGGAGGATTAAGGAAGACTTGCCTGATGAAC 966
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 967 ATTTTCCACATCAATACCTTGGTGTGATCAACTACTTCACTTCTCAACTTCTTAAA 1026
 321 GlySerSerGlnValIleAlaGlnTyrGluLeuIleArgLeuGlyHisGluGlyTyr 340
 1027 GGTTCATGACCCAAATGCTCAATATTTACCACTTATTCCTTGGTGGTGGGTTAC 1086
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 1087 AAGAATGTTATGGAGAAATTTGCAAGAAATGCAAGGTTATTAAGAGAAAGAAATGAA 1146
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 1147 AGTGGAAATTCACATAAATCCCAAAATTTGGAGTTCCTTTAGTAGCAATTTCTCT 1206
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTrp 400
 1207 AAGACAACAGTCAACACATGAGTTCGAAATTTCTGAACTTGTAGAAATTTGGATGG 1266
 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1267 ATTTGTTCTGGCATATACTATGCCCAAAATGCTCAACATGTCACAGTTCACAGATTGTC 1326
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1327 ATTAGAAGATTTCTCCCGCACACTAGCGGAGGACTGGTAAATAGACATTTGAAAAAGTC 1386
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460

1387 TTCCACGGAGTAGACACACTTCCGGCGAGGTCAACGCTAAGCTAGCCGTG----- 1437
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1438 ---GCCGAGGGAATGGCAGCGGCTGCAT-----AAGAAAACAGATAGAGAGTG 1485
 481 GlnArgAspIleIleThrGlyTyrLysLysPheValAlaAsp---ArgLysLysThrSer 499
 1486 CAGCTAGAGATTACTACTGCATGCTGAAATTTCTGCTGATAAAGAAAGAAAGACTAAT 1545
 500 GlyIleCys 502
 1546 GGAGTTTGT 1554

RESULT 6
 ABZ13132
 ID ABZ13132 standard; DNA; 1482 BP.
 XX
 AC ABZ13132;
 XX
 XX 21-JAN-2003 (first entry)
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 937.
 XX
 XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 XX WO200216655-A2.
 XX
 XX 28-FEB-2002.
 XX
 XX 24-AUG-2001; 2001WO-US26685.
 XX
 XX 24-AUG-2000; 2000US-227866P.
 XX
 XX 26-JAN-2001; 2001US-2664647P.
 XX
 XX 22-JUN-2001; 2001US-300111P.
 XX
 XX (SCRI) SCRIPES RES INST.
 XX
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 XX Harper JF, Kreps J, Wang X, Zhu T;
 XX
 XX WPI; 2002-304127/34.
 DR
 XX
 XX
 PT
 PT Identifying a stress condition to which a plant cell has been exposed
 and producing plants with increased tolerance to these abiotic stresses
 -
 XX
 PS Claim 144; SEQ ID NO 937; 577pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 cell has been exposed, comprising:
 CC (a) contacting nucleic acid representative of expressed polynucleotides
 in the plant cell with an array or probes representative of the plant
 cell genome; and
 CC (b) detecting a profile of expressed polynucleotides in the plant cell
 characteristic of a stress response. The method is useful in the
 production of transgenic plants, cells and seeds and in producing plants
 with increased tolerance to abiotic stress. The present sequence is that
 of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 in methods of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 specification but is based on sequence information supplied to Derwent by
 the European Patent Office.
 CC
 XX Sequence 1482 BP; 402 A; 331 C; 378 G; 371 T; 0 other;

Alignment Scores: 2.02e-212 Length: 1482
 Pred. No.: 2135.50 Matches: 411
 Score: 88.78% Conservative: 32
 Percent Similarity:

est Local Similarity: 82.36% Mismatches: 45
 uey Match: 81.66% Indels: 11
 B: 24 Gaps: 2
 S-10-006-852-2 (1-502) x ABZ13132 (1-1482)

1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 b 1 ATGGTTTTTGTCTAAGACAGATTCGGAATCTGATCTCAATCCATTCACATTTTGTCTCT 60
 Y 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 b 61 CGTTCAGTCCGCAACTCTCTCCACGATTCGAAATGCTCGAAGCTCGAAGCTCAATCCCAAAGAA 120
 Y 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 b 121 GCAGTTTACCAATCAATCAACAGCAGAGCTAATGCTCGATGTTAAACCAAGCTGAACCTA 180
 Y 61 AlaSerPheValThrThrProMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 b 181 GCTTCCCTTCGTGACCAATGGATGGAGCCAGAAATGTGACAAGCTCAATGATGGAGTCCATC 240
 Y 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 b 241 AACAGAACTACTGTCGATGCGAGTACCCCTGTCACCACTGAGCTTCAGAACCAGATGT 300
 Y 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 b 301 GTTAAACATGATAGCACCTCTCTCACCGCCGCTTGGTGAAGCTGAAAGCTGCGTGTGT 360
 Y 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 b 361 GTTGGCCCGTTCGATCGTCCGAGGCGGATATGTTGGCCGGTTCGCTTTTAAAGAGACA 420

141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 b 421 TGGCAGAATAAGCGTAAGGCCAAGGCTTCCCTTATGATAAGCCCAATATCGTAACCGGT 480
 Y 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 b 481 GCTAATGCTCCAGGTTTCTGGGAGAAATTCGCAAGGTAATTCGAAAGTGGAGCTTAAAGAA 540
 Y 181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 b 541 GTGAACCTAAGAGAAAGACTATACGTGATGGACCCCTGTAAGCGGCGFCGAAATGGTAGAC 600
 Y 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 b 601 GAAACACAAATTTGTGCTGCTGCCATCCTCGGTTCAACGTTAACCGGTGAATTCGAAGAC 660
 Y 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 b 661 GTTAAAGTCTCTCAACAGCTCTTGTGCGAGAAACAAAGCAAAACCCGGATGGGACGCGCA 720
 Y 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 b 721 ATACAGTGGACGAGGAGTGGTATGCTCCGTTCTTGTATCCGGAGCTGGAG 780
 Y 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 b 781 TGGGATTTCCGGCTACCGTTGGTTAAGAGTATTAATGTGAGTGTGCACAAATACGGTTTTG 840
 Y 281 ValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 b 841 GTTACGCGCGGTATGTTGGTGGTGTGAGAGAACCAAAACCCGATTTGCTGTGATGAACCT 900
 Y 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 b 901 ATCTTCCATATCAATATCTTGGCGCTGATCAACCAACTTTTACCTCAACTCTCCAAA 960
 Y 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 b 961 GGTTCAGTCAAGTATTGCTCAGTACTACCGAGCTGATTCGCTTGGATTTCGAGGGTTAT 1020

QY 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 Db 1021 CGCAATGTGATGGATAATTTTCGGGAAACATGATGGTACTAAGACAAGGATTAGAGAAA 1080
 QY 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 Db 1081 ACGGACGCTTTTAAATCGTCTCCAAAGAAAACCGGTGTCCTGTAGTGGCGTTTCTCTC 1140
 QY 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTyr 400
 Db 1141 AAGATAGTAGCCGCCACACGAGTTCGAGGTGCGCCATACACCTCCGTCGCTTGGCTGG 1200
 QY 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 Db 1201 ATCGTTCGGGCTACACGATGCTCGGATCGGAGCATGCTCACTGCTCTTCGAGTGT 1260
 QY 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 Db 1261 ATCCGAGAGATTTCTCGAACCTTACCCGAGAGATTTGGTAGCTGATTTCCGAGAAGTT 1320
 QY 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 Db 1321 CTACACGAGCTCGATACGCTTCCGGCGAGGTTCAAGCAAGATGGCTAATGGAAAAGTT 1380
 QY 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 Db 1381 AACGGT-----GTTAAGAGACGCCAGAGGAGACG 1410
 QY 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSer 499
 Db 1411 CAGAGAGAGTCACGGCTACTGGNAGAGTTGTFG---GAGACTAAGAGACCAAC 1464

RESULT 7
 AAL43411
 ID AAL43411 standard, cDNA, 1665 BP.
 XX
 AC AAL43411;
 XX
 DT 25-SEP-2002 (first entry)
 XX
 DE A thaliana GAD2 coding sequence.
 XX
 KW GAD; plant GABA production regulation; glutamic acid decarboxylase;
 XW Plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT CDS 17..1501
 FT /+tag= a
 FT /product= "GAD2"
 XX
 PN W0200238736-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 07-NOV-2001; 2001WO-0547447.
 XX
 PR 07-NOV-2000; 2000US-246367P.
 XX
 PA (EMER-) EMERALD BIOAGRICULTURE CORP.
 XX
 PI Kinnersley AM, Turano PJ;
 DR WPI; 2002-490073/52.
 DR P-PSDB; AAO15133.
 XX
 PT Making transformed plants that selectively increase gamma-aminobutyric
 PT acid production, by incorporating a DNA construct with a polynucleotide
 PT encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 XX

S Claim 18; Page 53-54; 63pp; English.
 X The present invention relates to a method of producing a transformed
 C plant that selectively increases production of gamma-aminobutyric acid
 C (GABA) in response to a signal, by incorporating into the plant's genome
 C a DNA construct with a non-constitutive promoter operably linked to a
 C polynucleotide encoding a functional plant glutamic acid decarboxylase
 C (GAD), to provide a transformed plant that expresses the GAD coding
 C sequence in response to a signal. Plants of this type have an enhanced
 C ability to tolerate environmental or other stresses. The present sequence
 C is the A. thaliana GAD2 coding sequence.
 X Sequence 1665 BP; 490 A; 311 C; 423 G; 441 T; 0 other;

Alignment Scores:
 red. No.: 8.88e-212 Length: 1665
 core: 2130.00 Matches: 404
 percent Similarity: 89.29% Conservative: 46
 est Local Similarity: 80.16% Mismatches: 42
 query Match: 81.45% Indels: 12
 B: 24 Gaps: 3

S-10-006-852-2 (1-502) x AAL43411 (1-1665)

Y 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 b 17 ATGGTTTGGACAAAACCCGCAACG--AATGATGAATCTGTCTGCACCAATGTTCCGATCT 73
 Y 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 b 74 CGCTATGTCGCACACTACCTCCCAAGTATGAGATGGTGGAGAAATCGATACCCGAAGAC 133
 Y 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 b 134 GCTGCATPACAGATCAATAAAGATGAGCTGATGCTTATGATGTTAAACCCGAGCTTAACCTA 193
 Y 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 b 194 GCTTCGTTTGTGACTACATGGATGGAAACCCAGAGTGTGACAAACTATCATGGAATCTATC 253
 Y 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 b 254 AACNAGAACTACGTTGATATGGATGAGTACCTGTGCACAACTGAGCTCCGAAACCCGATGT 313
 Y 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 b 314 GTAACATATTAGCTGACTGCTCAATCGCCACTCGAGAACTCGAGACGGCGGTGGGA 373
 Y 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 b 374 GTAGGGACAGTTGGTCTTCAGAAAGCCATCATGTTAGCCGGATTTGGCTTCAAAGAAA 433
 Y 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 b 434 TGGCAGAACAAACGCAAGGCTGAGGGTAAACCCCTATGACAAACCCCAACATTTGCTACTGGA 493
 Y 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 b 494 GCCAAATGTTCAAGTTTCTCGGAGAAATTCGCTCGTACTTCGAGGTGGAGCTAAAGGAA 553
 Y 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 b 554 GTAACCTAAGTGAAGTTACTACGTTGATGATCCAGACAAAGCAGCAAAATGGTAGAC 613
 Y 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 b 614 GAGAACCAATCTGTTCGCGACCAATATGGGATCCACACTCAACGGTGAAGTTCGAAGAC 673
 Y 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 b 674 GTGAAACCGTCTCAATGCTTGTAGTCAAGAAAACGAGGAGACTGGTTGGAACACACCG 733
 Y 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260

Db 734 ATCCACGTGGATGCAGCAAGTGGAGGTTTCATAGCTCCGTTTATCTATCTCCTGAATTAGAA 793
 Qy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 Db 794 TGGACTTTAGACTTCTTGGTTAAGAGTATCAACGTGAGTGTCAAGTATGGACTG 853
 Qy 281 ValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 Db 854 GCTATGCTGTTATGGTTGGTCTGTTGGAGGCGCAGCAGAGGATTTCCCTGAAGAGCTT 913
 Qy 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 Db 914 AFCFTTTCATATAAATATCTTGGTCTGATCAACCCACTTTCACCTCAATTTCTCCAAG 973
 Qy 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 Db 974 GGTTCGAGCCAAATTAATGCTCAATACTACCAGCTCATTCGCTTGGATTCGAGGGGTAC 1033
 Qy 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 Db 1034 AAAAAATGATGGAGAAATGTCATAGAGAAACATGGTGTCTTCAAAGAAAGGATAGAGAAA 1093
 Qy 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 Db 1094 ACAGAGCCTTTCACATAGTCTCAAAGGACCAAGGAGTCCAGTCCGCTTCTCTCTC 1153
 Qy 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
 Db 1154 AAGGACCATAGTTTCCACACAGAGTTCGAGATCTCTGAGATGCTACGTCGTTTGGCTGG 1213
 Qy 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 Db 1214 ATCFCCAGCTTACACTATGCTCCGCGGATGCACAGCACATCACGGTTCGCGGTGTC 1273
 Qy 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 Db 1274 AFCAGGGAGAGATTTCTCAAGAACACTCCGGAGAGACTTGTGCTGATATTTCCGAGGTG 1333
 Qy 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 Db 1334 CTTTCATGAGCTAGATACCTTCCCTTCCAAAGATATCTAAGAGATGGGAATA----- 1384
 Qy 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 Db 1385 -----GAAAGGATCCCGGAAATGTAAGAGGAGAAAGATGGGAGAG 1426
 Qy 481 Gln-----ArgAspIleIleIleThrGlyTrpLysLysPheValAlaAspArgLysThr 498
 Db 1427 GAGATTCGATGGAAGTATTGTTGGATGGAGAAAGTTTGTGAAGGAGAGGAGAAAGATG 1486
 Qy 499 SerGlyIleCys 502
 Db 1487 AATGGTGTGTGC 1498
 XX AAQ74691
 AC AAQ74691;
 XX 25-MAR-2003 (updated)
 DT 22-JUN-1995 (first entry)
 XX Early Ripening Tomato protein ERTD1 gene.
 XX Early Ripening Tomato; transgenic plant; crop improvement;
 KW fruit ripening; transformation; ss.
 XX Lycopersicon esculentum.
 XX W09421794-1.

RESULT 8

AAQ74691

ID AAQ74691 standard; cDNA; 1783 BP.

XX

AC AAQ74691;

XX

DT 25-MAR-2003 (updated)

DT 22-JUN-1995 (first entry)

XX

DE Early Ripening Tomato protein ERTD1 gene.

XX

XX Early Ripening Tomato; transgenic plant; crop improvement;

KW

fruit ripening; transformation; ss.

XX

OS Lycopersicon esculentum.

XX

PN W09421794-1.

XX

29-SEP-1994. 94WC-GB00581. 306 TGTGTTAAACATGTTAGCACATCTTTTCCAMGCCCGGTTGGTGTGATGATGAGACTGCAGTT 365
 22-MAR-1994; 93GB-0005859. 120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
 22-MAR-1993; 93GB-0005860. 366 GGAGTTGGTACAGTGGGTTTCATCAGAGCAATATGCTTGTGCTGCCTTGTCTTCAAGGC 425
 22-MAR-1993; 93GB-0005865. 140 LysTrpGlnAsnLysArgLysAlaGluLysProValAspLysProAsnIleValThr 159
 22-MAR-1993; 93GB-0005866. 426 AAATGGCAATCGAAAGAAAAGCAGAGGCAAACTTTCCGATAAGCCCTAATATAGTCACT 485
 22-MAR-1993; 93GB-0005868. 160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
 22-MAR-1993; 93GB-0005869. 486 GGAGCTAATGTGCAGTCTGCTGGAAAAATTTCAAGGTATTTGAGGTTGAGTTGAGG 545
 12-JUL-1993; 93GB-0014351. 180 GluValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetVal 199
 12-OCT-1993; 93GB-0020988. 546 GAGGTGAACCTAAAAGAGGATACTATGTAATGGACCCTGCCAAGCAGTAGAGATAGT 605
 (ZENECA LTD. 200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
 Abu-Bakar UK, Barton SL, Gallego-Veigas PP, Gray JE; 606 GATGAGAAATACAATATGTGTTGCTCAATCTCTGCTTCTACTCTGACTGGGAGTTGAG 665
 Grierson D, Lowe AL, Picton S, Whotton LC, Abubakar UK; 220 AspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
 WPI; 1994-317014/39. 666 GATGTAAGCTCTTAACGAGCTCTTACAAAAAGAAACAAGGAAACCCGGAITGGAGACA 725
 DNA constructs encoding fruit-ripening related proteins - useful 240 ProIleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
 for transformation of plants to modify fruit quality 726 CCGATTCATGTCGATGCTCGAGTGGAGGATTTATTGCTCTCTTCTCTGCCCAGATCTT 785
 Claim 1; Page 54; 74pp; English. 260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
 ERTD1 homologous transcripts are 1.8 kb in size. The transcript 786 GAATGGGATTTCCGTTTCCCTCTGTGAAAAAGTAAATGTCAGCGGFTCAAGATATGGC 845
 is found throughout early fruit development of both wild-type and 280 LeuValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeu 299
 rin (ripening inhibitor) fruit. The transcript is detected 846 CTTGTATATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 905
 throughout the entire ripening period in rin fruit. The transcript 300 LeuIlePheHisIleAsnTyrIleGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
 is not detected in leaves or wounded leaves of either rin or wild- 906 CTCGTCTTTCATATAACTCCTTGGTCTGATCAGCCTACTTTTACTCTCACTTCTCT 965
 type tomato. The sequence has been deposited as NCIMB 40598. The DNA 320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
 may be used to transform plants for production of plants with an altered 966 AAAGGTTCCATCAAAATAATGCCAGATTAATACAGTAAATAAGACTTTGGCTTGGAGGGT 1025
 C level of expression. Improvements in transgenic plants include 340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
 C resistance to damage and pests, longer shelf life, improved 1026 TATAGAGCTCATGAAGATTTGCTTATCAACCCAAAGTACTTACAGAGGAATCACA 1085
 C flavor/aroma, etc 360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
 C (Updated on 25-MAR-2003 to correct PN field.) 1086 AAAATGGGCGGTCGATATGCTCTAAGGATGTTGGGTGTTCTCTGTTGTAGCATTTTCT 1145
 X Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other; 380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
 Y Alignment Scores: 1146 CTCAGGACAGCAAAATATACGGTATTTGAGTATCTGAGCATCTCAGAAAGATTTGGA 1205
 b red. No.: 1.19e-199 Length: 1783 400 TrpIleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgVal 419
 core: 2014.00 Matches: 387 1206 TGGATCGTCCCTGCATACAAATGCCCGGATGCTGAACACATTTGCTGACTGGGGGT 1265
 percent Similarity: 86.44% Conservative: 53 420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
 est Local Similarity: 76.03% Mismatches: 55 1266 GTCATTAGAGAGGATTTCCAGCCACAGCCTAGCTGAGACTTGTCTGACATTTGAGAAA 1325
 very Match: 77.02% Indels: 14 440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle-----SerLeu 457
 B: 15 Gaps: 5 1326 ATTCTGTAGAGTTGGACACACAGCCTCTCTGTTGGCCCAAGCTGTCCTGCTGACT 1385
 S-10-006-852-2 (1-502) x AAQ74691 (1-1783) 458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeu-----MetValThrVal 473
 Y 1 MetValLeu---SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19 1386 GCTGAGGAGTCCGTGATCACAAAGGGTGTAGGGCTTCATCTTTTCATGATGGAFACTGTA 1445
 b 6 ATGGGTGTTAACACACCGCTCGATAAAGAGATTCAGAGAGACTTGCCTGTACATTTGCA 65
 Y 20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
 b 66 TCAAGATGATGTACAGGAACCTTTTACCTPAAGTTCAAAATGCTTAAATAATCCATGCCGAAA 125
 Y 40 GluAlaIleTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
 b 126 GAACAGCTTATCAGATGTYAAGCAGCAGACTTATGTTGGATGATGATGATGATGATGAT 185
 Y 60 LeuAlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
 b 186 TTAGCTTCTTTGTTAGCACATGGATGGCCGAGTGGCGATGATGATGATGATGATGATGAT 245
 Y 80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
 b 246 ATTAATAAAAATACTATGTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 305
 Y 100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaVal 119

D 29-SEP-1994. 94WC-GB00581. 306 TGTGTTAAACATGTTAGCACATCTTTTCCAMGCCCGGTTGGTGTGATGATGAGACTGCAGTT 365
 X 22-MAR-1994; 93GB-0005859. 120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
 F 22-MAR-1993; 93GB-0005860. 366 GGAGTTGGTACAGTGGGTTTCATCAGAGCAATATGCTTGTGCTGCCTTGTCTTCAAGGC 425
 R 22-MAR-1993; 93GB-0005865. 140 LysTrpGlnAsnLysArgLysAlaGluLysProValAspLysProAsnIleValThr 159
 R 22-MAR-1993; 93GB-0005866. 426 AAATGGCAATCGAAAGAAAAGCAGAGGCAAACTTTCCGATAAGCCCTAATATAGTCACT 485
 R 22-MAR-1993; 93GB-0005868. 160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
 R 22-MAR-1993; 93GB-0005869. 486 GGAGCTAATGTGCAGTCTGCTGGAAAAATTTCAAGGTATTTGAGGTTGAGTTGAGG 545
 R 12-JUL-1993; 93GB-0014351. 180 GluValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetVal 199
 R 12-OCT-1993; 93GB-0020988. 546 GAGGTGAACCTAAAAGAGGATACTATGTAATGGACCCTGCCAAGCAGTAGAGATAGT 605
 X (ZENECA LTD. 200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
 X Abu-Bakar UK, Barton SL, Gallego-Veigas PP, Gray JE; 606 GATGAGAAATACAATATGTGTTGCTCAATCTCTGCTTCTACTCTGACTGGGAGTTGAG 665
 X Grierson D, Lowe AL, Picton S, Whotton LC, Abubakar UK; 220 AspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
 X WPI; 1994-317014/39. 666 GATGTAAGCTCTTAAACGAGCTCTTACAAAAAGAAACAAGGAAACCCGGAITGGAGACA 725
 X DNA constructs encoding fruit-ripening related proteins - useful 240 ProIleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
 X for transformation of plants to modify fruit quality 726 CCGATTCATGTCGATGCTCGAGTGGAGGATTTATTGCTCTCTTCTCTGCCCAGATCTT 785
 X Claim 1; Page 54; 74pp; English. 260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
 X ERTD1 homologous transcripts are 1.8 kb in size. The transcript 786 GAATGGGATTTCCGTTTCCCTCTGTGAAAAAGTAAATGTCAGCGGFTCAAGATATGGC 845
 X is found throughout early fruit development of both wild-type and 280 LeuValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeu 299
 X rin (ripening inhibitor) fruit. The transcript is detected 846 CTTGTATATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 905
 X throughout the entire ripening period in rin fruit. The transcript 300 LeuIlePheHisIleAsnTyrIleGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
 X is not detected in leaves or wounded leaves of either rin or wild- 906 CTCGTCTTTCATATAACTCCTTGGTCTGATCAGCCTACTTTTACTCTCACTTCTCT 965
 X type tomato. The sequence has been deposited as NCIMB 40598. The DNA 320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
 X may be used to transform plants for production of plants with an altered 966 AAAGGTTCCATCAAAATAATGCCAGATTAATACAGTAAATAAGACTTTGGCTTGGAGGGT 1025
 X C level of expression. Improvements in transgenic plants include 340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
 X resistance to damage and pests, longer shelf life, improved 1026 TATAGAGCTCATGAAGATTTGCTTATCAACCCAAAGTACTTACAGAGGAATCACA 1085
 X flavor/aroma, etc 360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
 X (Updated on 25-MAR-2003 to correct PN field.) 1086 AAAATGGGCGGTCGATATGCTCTAAGGATGTTGGGTGTTCTCTGTTGTAGCATTTTCT 1145
 X Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other; 380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
 Y Alignment Scores: 1146 CTCAGGACAGCAAAATATACGGTATTTGAGTATCTGAGCATCTCAGAAAGATTTGGA 1205
 b red. No.: 1.19e-199 Length: 1783 400 TrpIleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgVal 419
 core: 2014.00 Matches: 387 1206 TGGATCGTCCCTGCATACAAATGCCCGGATGCTGAACACATTTGCTGACTGGGGGT 1265
 percent Similarity: 86.44% Conservative: 53 420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
 est Local Similarity: 76.03% Mismatches: 55 1266 GTCATTAGAGAGGATTTCCAGCCACAGCCTAGCTGAGACTTGTCTGACATTTGAGAAA 1325
 very Match: 77.02% Indels: 14 440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle-----SerLeu 457
 B: 15 Gaps: 5 1326 ATTCTGTAGAGTTGGACACACAGCCTCTCTGTTGGCCCAAGCTGTCCTGCTGACT 1385
 S-10-006-852-2 (1-502) x AAQ74691 (1-1783) 458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeu-----MetValThrVal 473
 Y 1 MetValLeu---SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19 1386 GCTGAGGAGTCCGTGATCACAAAGGGTGTAGGGCTTCATCTTTTCATGATGGAFACTGTA 1445
 b 6 ATGGGTGTTAACACACCGCTCGATAAAGAGATTCAGAGAGACTTGCCTGTACATTTGCA 65
 Y 20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
 b 66 TCAAGATGATGTACAGGAACCTTTTACCTPAAGTTCAAAATGCTTAAATAATCCATGCCGAAA 125
 Y 40 GluAlaIleTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
 b 126 GAACAGCTTATCAGATGTYAAGCAGCAGACTTATGTTGGATGATGATGATGATGATGAT 185
 Y 60 LeuAlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
 b 186 TTAGCTTCTTTGTTAGCACATGGATGGCCGAGTGGCGATGATGATGATGATGATGATGAT 245
 Y 80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
 b 246 ATTAATAAAAATACTATGTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 305
 Y 100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaVal 119

Y 474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAla 493
 b 1446 -----GAGACTCAGAAAAGACATTATCAAAACATTGGAGAAAATCGCAGGG 1490
 Y 494 AspArgLysThrSerGlyIleCys 502
 b 1491 -----AAGAAGACCAGCGGAGTCTGC 1511
 RESULT 9
 AAL43418
 D AAL43418 standard; cDNA; 1783 BP.
 X
 C AAL43418;
 T 25-SEP-2002 (first entry)
 T Tomato GAD coding sequence.
 E
 X GAD; plant GABA production regulation; glutamic acid decarboxylase;
 W plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
 X
 X Lycopersicon esculentum.
 X
 H Location/Qualifiers
 I CDS 6..1514
 T /*tag= a
 T /product= "GAD"
 X W0200238736-A2.
 X
 D 16-MAY-2002.
 X
 X 07-NOV-2001; 2001WO-US47447.
 F
 X 07-NOV-2000; 2000US-246367P.
 R
 X (EMER-) EMERALD BIOAGRICULTURE CORP.
 X
 X Kinnersley AM, Turano FU;
 I
 X WPI; 2002-490073/52.
 R
 X P-PSDB; AAO15140.
 X
 T Making transformed plants that selectively increase gamma-aminobutyric
 T acid production, by incorporating a DNA construct with a polynucleotide
 T encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 T
 X Claim 18; Page 59; 63pp; English.
 S
 X The present invention relates to a method of producing a transformed
 C plant that selectively increases production of gamma-aminobutyric acid
 C (GABA) in response to a signal, by incorporating into the plant's genome
 C a DNA construct with a non-constitutive promoter operably linked to a
 C polynucleotide encoding a functional plant glutamic acid decarboxylase
 C (GAD), to provide a transformed plant that expresses the GAD coding
 C sequence in response to a signal. Plants of this type have an enhanced
 C ability to tolerate environmental or other stresses. The present sequence
 C is the tomato GAD coding sequence.
 X
 Q Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other;

Alignment Scores:
 red. No.: 1.19e-199 Length: 1783
 core: 2014.00 Matches: 387
 Percent Similarity: 86.44% Conservative: 53
 est Local Similarity: 76.03% Mismatches: 55
 very Match: 77.02% Indels: 14
 B: 24 Gaps: 5


```

b 1086 AAAATGGCGGTCGATATGCTTAAGGATGGGGTTCCTGTGTAGCATTTCT 1145
Y 380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGly 399
b 1146 CTCAGGACAGACAAATATACGGTATTGAGATCTGAGCACTCAGAAATTGGA 1205
Y 400 TrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgVal 419
b 1206 TGGATCGTCCCTGCATACACAATGCCACCGGATGCTGAACACATGCTGTACTGGCGGTT 1265
Y 420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
b 1266 GTCAATAGAGAGGATTCAGCCACACAGCTAGCTGAGAGACTGTCTTTCACATGAGAAA 1325
Y 440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisIlysIle-----SerLeu 457
b 1326 ATTCTGTGAGAGTTGGACACACAGCTCTCTCGTTTCCACCACCAAGCTCCCGTGTCACT 1385
Y 458 GlyGlnGluLysSerGluSerAsnSerAsnLeu-----MetValThrVal 473
b 1386 GCTGAGGAACTGGCTGATGACAAAGGATGATGGGCTTCATCATTTTCACATGGATACTGTA 1445
Y 474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAla 493
b 1446 -----GAGACTCAGAAGACATTAACAATTTGGAGGAAAATCGCAGGG 1490
Y 494 AspArgLysLysThrSerGlyIleCys 502
b 1491 -----AAGAAGACCAGCGGACTGTC 1511
RESULT 10
AAC36324
AAC36324 standard; DNA; 1605 BP.
X D X 17-OCT-2000 (first entry)
X E X Arabidopsis thaliana DNA fragment SEQ ID NO: 13376.
X W X Hybridisation assay; genetic mapping; gene expression control;
X W X protein identification; signal transduction pathway;
X W X metabolic pathway; promoter; termination sequence; ss.
X S X Arabidopsis thaliana.
X N X EP1033405-A2.
X D X 06-SEP-2000.
X F X 25-FEB-2000; 2000EP-0301439.
X R X 25-FEB-1999; 99US-0121825.
X R X 03-MAR-1999; 99US-0123180.
X R X 03-MAR-1999; 99US-0123548.
X R X 23-MAR-1999; 99US-0125788.
X R X 25-MAR-1999; 99US-0126264.
X R X 29-MAR-1999; 99US-0126785.
X R X 01-APR-1999; 99US-0127462.
X R X 06-APR-1999; 99US-0128234.
X R X 08-APR-1999; 99US-0128714.
X R X 16-APR-1999; 99US-0129845.
X R X 19-APR-1999; 99US-0130077.
X R X 21-APR-1999; 99US-0130449.
X R X 23-APR-1999; 99US-0130510.
X R X 23-APR-1999; 99US-0130891.
X R X 28-APR-1999; 99US-0131449.
X R X 30-APR-1999; 99US-0132048.
X R X 30-APR-1999; 99US-0132407.
X R X 04-MAY-1999; 99US-0132484.
X R X 05-MAY-1999; 99US-0132485.
X R X 06-MAY-1999; 99US-0132486.

```

```

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

```

```

R R R 22-JUL-1999; 99US-0145192.
R R R 23-JUL-1999; 99US-0145145.
R R R 23-JUL-1999; 99US-0145218.
R R R 23-JUL-1999; 99US-0145224.
R R R 26-JUL-1999; 99US-0145276.
R R R 27-JUL-1999; 99US-0145913.
R R R 27-JUL-1999; 99US-0145918.
R R R 27-JUL-1999; 99US-0145919.
R R R 28-JUL-1999; 99US-0145951.
R R R 02-AUG-1999; 99US-0146386.
R R R 02-AUG-1999; 99US-0146388.
R R R 03-AUG-1999; 99US-0146389.
R R R 03-AUG-1999; 99US-0147038.
R R R 04-AUG-1999; 99US-0147204.
R R R 04-AUG-1999; 99US-0147302.
R R R 05-AUG-1999; 99US-0147192.
R R R 06-AUG-1999; 99US-0147260.
R R R 06-AUG-1999; 99US-0147303.
R R R 06-AUG-1999; 99US-0147416.
R R R 09-AUG-1999; 99US-0147493.
R R R 09-AUG-1999; 99US-0147935.
R R R 10-AUG-1999; 99US-0148171.
R R R 11-AUG-1999; 99US-0148319.
R R R 12-AUG-1999; 99US-0148341.
R R R 13-AUG-1999; 99US-0148565.
R R R 13-AUG-1999; 99US-0148684.
R R R 16-AUG-1999; 99US-0149368.
R R R 17-AUG-1999; 99US-0149175.
R R R 18-AUG-1999; 99US-0149426.
R R R 20-AUG-1999; 99US-0149722.
R R R 20-AUG-1999; 99US-0149723.
R R R 20-AUG-1999; 99US-0149829.
R R R 23-AUG-1999; 99US-0149902.
R R R 25-AUG-1999; 99US-0149930.
R R R 25-AUG-1999; 99US-0150566.
R R R 26-AUG-1999; 99US-0150884.
R R R 27-AUG-1999; 99US-0151065.
R R R 27-AUG-1999; 99US-0151066.
R R R 27-AUG-1999; 99US-0151080.
R R R 30-AUG-1999; 99US-0151303.
R R R 31-AUG-1999; 99US-0151438.
R R R 16-SEP-1999; 99US-0151930.
R R R 17-SEP-1999; 99US-0152363.
R R R 17-SEP-1999; 99US-0153070.
R R R 13-SEP-1999; 99US-0153758.
R R R 15-SEP-1999; 99US-0154018.
R R R 16-SEP-1999; 99US-0154039.
R R R 20-SEP-1999; 99US-0154779.
R R R 22-SEP-1999; 99US-0155139.
R R R 23-SEP-1999; 99US-0155486.
R R R 24-SEP-1999; 99US-0155659.
R R R 28-SEP-1999; 99US-0156458.
R R R 29-SEP-1999; 99US-0156596.
R R R 04-OCT-1999; 99US-0157117.
R R R 05-OCT-1999; 99US-0157753.
R R R 06-OCT-1999; 99US-0157865.
R R R 07-OCT-1999; 99US-0158029.
R R R 08-OCT-1999; 99US-0158232.
R R R 12-OCT-1999; 99US-0158369.
R R R 13-OCT-1999; 99US-0159293.
R R R 13-OCT-1999; 99US-0159294.
R R R 13-OCT-1999; 99US-0159295.
R R R 14-OCT-1999; 99US-0159329.
R R R 14-OCT-1999; 99US-0159330.
R R R 14-OCT-1999; 99US-0159331.
R R R 14-OCT-1999; 99US-0159637.
R R R 14-OCT-1999; 99US-0159638.
R R R 18-OCT-1999; 99US-0159584.
R R R 21-OCT-1999; 99US-0160741.
R R R 21-OCT-1999; 99US-0160767.
R R R 21-OCT-1999; 99US-0160768.
R R R 21-OCT-1999; 99US-0160770.
R R R 21-OCT-1999; 99US-0160814.

```

```

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1.3e-194 Length: 1605
Score: 1965.00 Matches: 372
Percent Similarity: 86.29% Conservative: 56
Best Local Similarity: 75.00% Mismatches: 62
Query Match: 75.14% Indels: 6
DB: 21 Gaps: 3

US-10-006-852-2 (1-502) x AAC36324 (1-1605)

Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 40 ANGTFACFCGCA--ACCAACTCTGACTCCGACGAGCATTTTGCATTTCCATTTTGCCTTCT 96
Qy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 97 AGATATGTCGGTGTGTGTTCCAGGTTCAAGATGCTGACCATTCATGCCAARAGAT 156
Qy 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 157 GCTGCTTATCAAGTATCAATGATGATGATGCTTGTATGATGATGATGATGATGATGATGAT 216
Qy 61 AlaSerPheValThrThrTriMetGluProGluCysAspLysLeuIleMetSerSerIle 80
Db 217 GCCTCTTTTGCACCCCTTGGATGGAACCTGAGTGTGACAAACTCATCGATTTCTGTC 276
Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
Db 277 ANPAGAACTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 336
Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
Db 337 GTAAATATGATAGCAAACTTTTCCATGCTCCGCTGGAGAAAGACGAGGCTGCTATTGGG 396
Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
Db 397 TGTGAACTGTGGTTCATCTGAGGCTATAAATGCTTGTCTGTTTGGCTTTTCAAAGGANA 456
Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
Db 457 TGGCAACATAGAGAAAAGCTCAGGGTCTACTATTGATAAGCCCTAACATTGTCACCTGA 516
Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
Db 517 GCCAATGTTCCAGTGTGCTGGGAGAAGTTTCAAGAGTACTTTGAGGTAGAGCTCAAAGAG 576
Qy 181 VallysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
Db 577 GTGAACTAAGTGAAGACTACTATGTTATGGATCCAGTAAAGCTGTAGAGATGGTGGAT 636
Qy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
Db 637 GAGAATACCATCTGTGTTGCAGCAATCTAGSATCTACACTTACTGGAGAGTTTGAGGAC 696
Qy 221 VallysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
Db 697 GTTAAAGCAATTGAACGATCTCTTAGCTGAGAAAAAACCAGAGACAGGATGGGAAACTCT 756

```

241 IleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 |||||
 757 ATTATGTTGATGCAGCCAGTGGAGATTCAITGCTCCCTTACCTGATCTTGAA 816
 |||||
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 |||||
 817 TGGGACTTTAGCTTCCATGGGTGAAGAGTATTAACGTCAGTGGTGCACAAATGGACIT 876
 |||||
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeu 300
 |||||
 877 GTGTATGCAGGAGTTGGTTGGTTGTCGGAGAACAAAGATGATTTGCCAGGAACTT 936
 |||||
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 |||||
 937 GTCTCCCATCACTAATCTGGGAGCTGATCAACCCTTCTCACTCAACTTCTCAAAA 996
 |||||
 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 |||||
 997 GGGTCGAGCCAAATCACTGCTAGTACTATCACTTTATCCGACTAGCTTTGAGGGATAC 1056
 |||||
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlyLys 360
 |||||
 1057 AAGACATAATGAAACTGCATGGATAACGCCAAGGAGGCTAAGAGAGAAATAGAGATG 1116
 |||||
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 |||||
 1117 ACAGGAAGTTCAACTTGTCTCAAGATATTGGCGTGCCACTAGTGCATTTCTCTCTC 1176
 |||||
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTyr 400
 |||||
 1177 AAAGACAGTAGCAAGCACACGGTGTGGATCGCAGAGTCTTTGAGAAATTCGGGTGG 1236
 |||||
 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 |||||
 1237 ATCATACCGGCTTACACTATATGCTCCAGATGCACAGCAGCATTTGCTGTCTCAGAGTTGG 1296
 |||||
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGlyLysVal 440
 |||||
 1297 ATAAGAGAAGACTTTTATGCGGAGGCTTGCAGATAGACTCATCACATATCATTCAGGTTG 1356
 |||||
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 |||||
 1357 CTGAAGAGATTGAAGGGCTTCTAGCAGATTGCACATCTTCTGCTGCG----- 1404
 |||||
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 |||||
 1405 GCTGCAGCGTTAGTGTGATGATGAAGAATTAAAGTGAAGACTGCCAAGATGCTCTTG 1464
 |||||
 481 GlnArgAspIleIleThrGlyTyrLysPheValAlaAspArgLys 496
 |||||
 1465 GAG--GATATCACTAAGTATTGGAAACCGCTTGTGGAAACACAAAGAGA 1509
 |||||

RESULT 11
 AL43413
 D AAL43413 standard; DNA; 2121 BP.
 X C
 X C AAL43413;
 X T
 X T 25-SEP-2002 (first entry)
 X E A thaliana GAD4 coding sequence.
 X E GAD; plant GABA production regulation; glutamic acid decarboxylase;
 X W plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
 X X Arabidopsis thaliana.
 X S
 X N W0200238736-A2.
 X N
 X X 16-MAY-2002.
 X D
 X D 07-NOV-2001; 2001WO-US47447.
 X F
 X X

07-NOV-2000; 2000US-246367P.
 (EMER-) EMERALD BIOAGRICULTURE CORP.
 Kinnersley AM, Turano FJ;
 WPI; 2002-490073/52.
 P-PSDB; AAO15135.
 Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 18; Page 55; 63pp; English.
 The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the A. thaliana GAD4 coding sequence.
 Sequence 2121 BP; 626 A; 412 C; 473 G; 610 T; 0 other;

Alignment Scores:
 Pred. No.: 1.8e-191 Length: 2121
 Score: 1936.50 Matches: 410
 Percent Similarity: 61.99% Conservative: 32
 Best Local Similarity: 57.50% Mismatches: 46
 Query Match: 74.05% Indels: 226
 DB: 24 Gaps: 7
 US-10-006-852-2 (1-502) x AAL43413 (1-2121)
 QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 1 ATGGTTTTGTCTAAGACAGTTCCGAACTGATGCTCAATCCATTCACCTTTTGCCTTCT 60
 QY 21 ArgTyrValArgThrSerLeuPro----- 28
 Db 61 CGTACGTCGCGAACTCTCTCCACGGTAAACAATGTPAACAAAATCTTTTGTCTAATG 120
 QY 28 ----- 28
 Db 121 TTTTCGTCACAAATAGTAAACATGTAATGATCTAAACCTTGGATAGCTTTTTTTTGGCGC 180
 QY 28 ----- 28
 Db 181 TGGTTAAATGTTGTAGATTTATATGTTTATATATACTATAAGGAAGGACATGTTTCGTTAT 240
 QY 29 -----ArgPheLysMetProGluAsnSerIlePro 38
 Db 241 TTTAACTTAATGATATCATCATTTTCAATTTAGATTGCGAAATGCGCTGAGAACTCAATCCCA 300
 QY 39 LysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeu 58
 Db 301 AAAGAAGCAGCTTACCAAATCATCAACGCGAGCTAATGCTCGATGGTAAACCCAAAGGCTG 360
 QY 59 AsnLeuAlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSer 78
 Db 361 AACCTAGTTCCTTCGTGACCACATGGATGGAGCCAGAAATGTGACAACTCATGATGGAG 420
 QY 79 SerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsn 98
 Db 421 TCCATCAACAAGAACTACGTGCATGGACGAGTACCCTGTCACCACCTGAGCTTCAGAAC 480
 QY 99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAla 118
 Db 481 CGATGTGTTAACATGATAGCAGCTCTCTTCAACGGCGGCTTGGTGAGGGTGAAGCTGCC 540
 |||||

119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
 541 GTTGGTGTGGCACCGTCGGATCGTCGGAGCGGATATGTTGGCCGGTTTGGCTTTTAAAG 600
 139 ArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleVal 158
 601 AGACAATGCCAGAAATAGCGTAAGGCCCAAGGGCTTCCCTTATGATAAGCCCAATATCGTA 660
 159 ThrGlyAlaAsnVal-----163
 661 ACCGGTCTAATGTCCAGGTAAACCAAAACAATAATGATGAAATATTAAACAAGACAAA 720
 164 -----GlnVa 165
 721 ATTGAATTTTCAATCCGGTTAAGTTATATGTTGACTCAATTTCCGGTTCAATACAGGT 780
 165 lCysTrpGluLysPheAlaArgTy:PheGluValGluLysLysLysLysLeuSerG1 185
 781 TTGCTGGGAGAAATTCGCAAGGTATTTCCAAAGTGGAGCTTAAGGAAGTGAACCTAAGAGA 840
 185 uGlyTyTrpValMetAspProGlnAlaValAspMetValAspGluAsnThrIleCy 205
 841 AGACTATYACGTGTGGACCCCTGTAAAGCGGTGAAATGGTAGACGAAACACAAATTTG 900
 205 sValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspValLysLeuLeuAs 225
 901 TGTCCGTGCCATCTCCGTTCACCGTTAAACCGTTAAACCGTTAAACCGTTAAACCGTTCAA 960
 225 rAspLeuLeuValGluLysAsnLysGluThiGly-----236
 961 CGACCTCTTGTGGAGAAAACAAGCAAAACCGG- GTAATAAACAACCAACCGGAGAAACAAG 1019
 236 -----236
 1020 CTAATATCGAATGTAATCGGTTGGAGTCCGGTTTTAAGCTTCAAACAACAAATTTGCAG 1079
 237 -TrpAspThrProIleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTy 256
 1080 ATGGGACACGCCAATACACGTGGAGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1139
 256 rProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHi 276
 1140 TCCGGAGCTGGAGTGGGATTTCCGGCTACCGTTGGTTAAGATATTAATGTGAGTGGTCA 1199
 276 sLysTyGlyLeuValTyAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLe 296
 1200 CAAAATACGGTTGGTTTACCGCGTATTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTT 1259
 296 uProGluGluLeuIlePheHisIleAsnTyziLeuGlyAlaAspGlnProThrPheThrie 316
 1260 GCCTGATGAACCTATCTCCATATCAATTTCTTGGCGGTGATCAACCAACCGTTTACACT 1319
 316 uAsnPheSerLys-----320
 1320 CAACCTTCTCCAAAGGTACATTACCATAAGTCCATACATATAATAATTTCATATAATTT 1379
 321 -----Glyse 322
 1380 TTGGTGTATGAAATGTTTTATAGACTAAACATTTGATAAATCGTTGTATAAACCGGTTTC 1439
 322 rSerGlnValIleAlaGlnTyTyGlnLeuIleArgLeuGlyHisGlu-----338
 1440 AAGTCAAGTGAATTCCTCAGTACTACCAGCTGATTCCTTGGATTTCGAGGTAATAATAA 1499
 338 -----338
 1500 CTCAAAATAGCAATATATTACCAAAATGGTCAATAAAGAAACTAGATGATATATATTTA 1559
 338 -----338
 1560 AGTTGTTACTTGTACTATACCTTTGAAATTAACCGTTCCCTAACATGACTAGTTTGGTATT 1619

339 -----GlyTyArgAsnValMetGluAsnCy 347
 1620 GTGTAATTAATAATGTTTTCTTGTGTTGATTTAGGGTTATCGCAATGTGATGATAAATTTG 1679
 347 sArgGluAsnMetIleValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleVa 367
 1680 TCGGGAAAACATGATGGTACTAAGCAAGATAGAGAAAACGGGACGTTTTAAAAATCGT 1739
 367 lSerLysAspGluGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHisTh 387
 1740 CTCCAAAGAAAACGGTGTCCGTTAGTGGCGTTTTCTCAAAGATAGTAGCCGCCACAA 1799
 387 rGluPheGluIleSerAspMetLeuArgTyGlyTrpIleValProAlaTyTrpThrMe 407
 1800 CGAGTTCGAGGTGGCCATACACTCCGTCGCTTCGGTGGATCGTTCCGGCTTACACGAT 1859
 407 rProProAsnAlaGlnHisIleThrValLeuArgValIleAtrGluAspPheSerAr 427
 1860 GCCTGGGATGCCAGCAIGTCACTGCTCTCCAGTTGTTATCCGAGAGATTCTCTCG 1919
 427 gThrLeuAlaGluArgLeuValIleAspIleGluLysValMetArgGluLeuAspGluLe 447
 1920 AACCTTAGCCGAGAGATGGTAGCTGATTTTCGAGAAGGTTCTACACGAGCTCGATACGCT 1979
 447 uProSerArgValIleHisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSerAs 467
 1980 TCCGGCGAGGGTTTCCGCCCAAGATGGCTAATGGAAAAGTTAAACGGT-----2025
 467 pAsnLeuMetValThrValLysLysSerAspIleAspLysGlnArgAspIleIleThrG1 487
 2026 -----GTTAAGAGAGCCGACGAGAGAGCCGAGAGAGAGAGAGAGAGTCAAGCCCTA 2069
 487 yTrpLysLysPheValAlaAspArgLysLysThrSer 499
 2070 CTGGAGAAAGTTGTG--GAGACTAAGAAAGCAAC 2103
 RESULT 12
 AA170004
 ID AA170004 standard; DNA; 1180 BP.
 XX AA170004;
 AC AA170004;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Tomato glutamate decarboxylase clone T-gad-19 DNA sequence.
 KW
 XX
 OS Lycopersicon esculentum.
 XX
 FN JP2001238555-A.
 XX
 PD 04-SEP-2001.
 XX
 PF 13-DEC-2000; 2000JP-0404172.
 XX
 PR 13-DEC-1999; 99JP-0352552.
 XX
 PA (AJIN) AJINOMOTO KK.
 XX
 DR WPI; 2001-610050/70.
 XX
 PT Creation of a transgenic plant of improved amino acid composition,
 XX involves transformation with antisense sequences -
 PS Example 1; Page 13-14; 18pp; Japanese.
 XX
 CC The present invention relates to a method for creating a transgenic plant
 CC which accumulates a larger amount of at least one amino acid of edible
 CC part than a natural plant of the same species cultured under a same
 CC condition. The method involves transforming a plant with a genetic
 CC structure containing an antisense sequence of a gene sequence encoding
 CC glutamate decarboxylase (GAD) and a controlling sequence which can

C express the antisense sequence. The resultant transgenic plant has high
 C added value. The present sequence is the DNA sequence for clone T-gad-19
 C of the tomato GAD gene, which was used to illustrate the present
 C invention.

X Q Sequence 1180 BP; 372 A; 188 C; 276 G; 344 T; 0 other;

Alignment Scores:

red. No.:	2,32e-181	Length:	1180
core:	1836.00	Matches:	347
ercent similarity:	94.40%	Conservative:	24
est Local Similarity:	88.30%	Mismatches:	22
uery Match:	70.21%	Indels:	1
B:	22	Gaps:	0

IS-10-006-852-2 (1-502) x RAI70004 (1-1180)

Y	38	ProLysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArg	57
b	1	CCAAAGGAAGCAGCATATCAGATTCTAATGATGAACTTATGTTAGATGAAATCCAAGG	60
Y	58	LeuAsnLeuAlaSerPheValThrTirMetGluProGluCysAspLysLeuIleMet	77
b	61	TTGAAATTTGGCACTCTTTTGTGCAACAATGGATGGACCAAGATGTGACAAATTTGATGATG	120
Y	78	SerSerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln	97
b	121	GATTCATTAACAAAATATGTTGACATGGATGAATATCTCTCACCACCTGAGCTTCAG	180
Y	98	AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr	117
b	181	AATCGGTGTGAAACATGATAGCCGATTTATTAATGCACCACCTTGAAGATGGAGAAACT	240
Y	118	AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe	137
b	241	GCAGTTGGATGGAAACAGTTGGTCTTCAGAGCCATATGCTTGTGGATTTGGCCTTT	300
Y	138	LysArgLysTrpGlnAsnLysArgLysAlaGluCysProValAspLysProAsnIle	157
b	301	AAGAAATAATGGCAAAACAAATAAGAAAGCCCAAGAAAGCCCTATGATAGCCCAACT	360
Y	158	ValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGlu	177
b	361	GTTACTGTGCTAATGTCAGGTGTGGGAAAATTTGCAAGGTAATTTTGAAGTTGAG	420
Y	178	LeuLysGluValLysLeuSerGluClyTyrTyrValMetAspProGlnGlnAlaValAsp	197
b	421	CTAAAAGAAAGTGAAGTTGACTGATGATGATGATGATGATGATGATGATGATGATGATG	480
Y	198	MetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGlu	217
b	481	ATGGTGTGATGAGAACACAAATTTGTAGTCTCTAATTTTGGGTTCAACTCTGAATGGGGAG	540
Y	218	PheGluAspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrp	237
b	541	TTTGAGGAAGTGAAGAAATTTGAATGACCTCTCTATT-GAAAAGAACAGGAAACAGGGTGG	599
Y	238	AspThrProIleHisValAspAlaSerGlyPheIleAlaProPheLeuTyrPro	257
b	600	GACACTCCAATTCATGTGGATGCAGCTAGTGTGGATTTATTTGGGTTCAACTCTGAATGGGGAG	659
Y	258	GluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLys	277
b	660	GAACTTGAATGGACCTTTAGATTGCCATTGTTAAAGATATAAATGTCAGCGGTCACAG	719
Y	278	TyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuPro	297
b	720	TATGGCTTGTATATGCTGGTGTGGTTCATATGTCGGGAGCAAGAAAGACTTTGCC	779
Y	298	GluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsn	317
b	780	GATGAACCTTAATTTTTCATATTAATTTACCTTGGGTCTGATCAACCTACTTTCCACCTCAAT	839

OY	318	PheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleLeuLeuGlyHis	337
Db	840	TTCTCTAAAGGATCAAGTCAAGTAATGTCICAATATTAATCAACTAATTCGTTTGGTTAT	899
OY	338	GluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGly	357
Db	900	GAGGTTTATAAGAACGTCATGAAGATTTGCTTATCAAAACGCAAGGCTACTAACAGAGGA	959
OY	358	LeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAla	377
Db	960	ATCACAATAATGGGAAGATTCGAGTTTGTCTCCAAAGGATGTTGGTTCCTTGGTTCGA	1019
OY	378	PheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArg	397
Db	1020	TTTTCTCTTAAAGACAATAGCAAAACATGATGAGTTTGAAGTATCTGAAACTTTAAAGGAGA	1079
OY	398	TyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeu	417
Db	1080	TTTGTATGATGTTTCCAGCATACACATATGCCAGATGCCAGATGCTCAACACATCACAGTTC	1139
OY	418	ArgValValIleArgGluAspPheSerArgThrLeuAla	430
Db	1140	AGAGTTGTTATTAGAGAAGATTTCTCCCGTACACTCGCC	1178
RESULT 13			
AAAL43412			
ID	AAAL43412	standard; DNA;	2493 BP.
XX			
AC	AAAL43412;		
XX			
DT	25-SEP-2002	(first entry)	
XX			
DE	A thaliana GAD3 coding sequence.		
XX			
KW	GAD; plant GABA production regulation; glutamic acid decarboxylase;		
XX			
KW	plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	WO200238736-A2.		
XX			
XX	16-MAY-2002.		
XX			
PF	07-NOV-2001;	2001WO-US47447.	
XX			
PR	07-NOV-2000;	2000US-246367P.	
XX			
XX	(EMER-) EMERALD BIOAGRICULTURE CORP.		
XX			
PI	Kimmersley AM,	Turano FU;	
XX			
DR	WPI;	2002-490073/52.	
DR	P-PSDB;	AAO15134.	
XX			
XX	Making transformed plants that selectively increase gamma-aminobutyric		
PT	acid production, by incorporating a DNA construct with a polynucleotide		
PT	encoding a plant glutamic acid decarboxylase enzyme into plant's genome		
PT			
XX			
PS	Claim 18; Page 54-55; 63pp; English.		
XX			
CC	The present invention relates to a method of producing a transformed		
CC	plant that selectively increases production of gamma-aminobutyric acid		
CC	(GABA) in response to a signal, by incorporating into the plant's genome		
CC	a DNA construct with a non-constitutive promoter operably linked to a		
CC	polynucleotide encoding a functional plant glutamic acid decarboxylase		
CC	(GAD), to provide a transformed plant that expresses the GAD coding		
CC	sequence in response to a signal. Plants of this type have an enhanced		
CC	ability to tolerate environmental or other stresses. The present sequence		
CC	is the A. thaliana GAD3 coding sequence.		
XX			
SQ	Sequence 2493 BP; 783 A; 458 C; 507 G; 745 T; 0 other;		

QY	194	GlnAlaValAspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThr	213
Db	960	AAAGCGGTTGAAATGGTAGAGCAAAACACATATATGCGTCTGGGCCAATCCCTCGGTCGACA	1019
QY	214	LeuAsnGlyGluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLys	233
Db	1020	CTAACCGGAGAAATCGAAGACGTTAAGCTCCTCAACGACCTTTTAGTCGAGAAAAACAAG	1079
QY	234	GluThrGly	236
Db	1080	AAACCGGGTAATTAATCAAAACCACTAACAAATTAATTTATATACCTTTTCCTGCTAGA	1139
QY	236		236
Db	1140	AATATTACAATTTCTAAACGTCAGATATATTTGCTTTAGAAAATTTTTATTTTTTGAATGAA	1199
QY	236		236
Db	1200	TATAAACTTATTAAACCAAAACAAACCAATATATGTTTACATTTATATGCTTCCTTTGATC	1259
QY	236		236
Db	1260	GAATGGTCTTTTAAATACTGATTAATAAAATGTTTTGCTTAAATAATTAACAATTTAAT	1319
QY	236		236
Db	1320	GTGAGATATTCAAGCAATTTCTAATATCAAAACCGATAAAACAAACAACTGATTTAAT	1379
QY	237		243
Db	1380	TATTTAACCGGTTTGGTTCGGTTTAAATATATTTGTAGATGGGATACCGCAATTCAGTG	1439
QY	244	AspAlaIaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTyrAspPhe	263
Db	1440	GACGACGAGTGGTGGTATTGCTCCCTTCTGTTATCGGACTTGGAGTGGATTC	1499
QY	264	ArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeuValTyrAla	283
Db	1500	CGTATCCGTTGGTTAAAGCATAAATGTGAGTGTCAAAATACGGTTCGGTTGGTTAGCC	1559
QY	284	GlyIleGlyTyrValIleTyrArgAsnLysGluAspLeuProGluLeuIlePheHis	303
Db	1560	GGTATCCGTTGGTGGTATGAGAAACCAAAACCGATTTGCCTGATGAACTTATCTCCAT	1619
QY	304	IleAsnTyrLeuGlyValAlaAspGlnProThrPheThrLeuAsnPheSerLys	320
Db	1620	ATCAATATCTTGGAGCTGATCAACCCACATTTACCCCTCAACTTCTTAAGGTACATTA	1679
QY	320		320
Db	1680	CCATACTTATGTAAGTTTAGATATATTTATAGATTAATGTTTTGTTAATCTTTGTATA	1739
QY	321		338
Db	1740	TTACCAGGTCAGTCRAGTGGTTCCTCGTACCTGATGAACTTCTTAAGGTACGAG	1799
QY	338		338
Db	1800	GTAATAATAACTCAATAAAGAACTAAACCGTTACTATAAATCCAAATCGTATACGTACTAG	1859
QY	338		338
Db	1860	TATAATACAAAGTTGTTACTATATCTTTATGACTACAAAAGTTTCAAAAACCAAGAATGTC	1919
QY	338		338
Db	1920	TAAATACATCCATAAGATTAAACGTTCCCTTAAATTTGACAAGTTTTGTTGTAGATAG	1979
QY	339		350
Db	1980	CTAATAATCTTTTGGTTGGTTAGGATATCGCAACGATGATGATTAATTCCTCCCGGAA	2039

Y	1	MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer	20
b <td>1 <td>ALGGITTTTACTAAGACAGCTCCAAATCCGATGATTCATCCATTCACATTTTGGCTCC <td>60</td> </td></td>	1 <td>ALGGITTTTACTAAGACAGCTCCAAATCCGATGATTCATCCATTCACATTTTGGCTCC <td>60</td> </td>	ALGGITTTTACTAAGACAGCTCCAAATCCGATGATTCATCCATTCACATTTTGGCTCC <td>60</td>	60
Y <td>21 <td>ArgTyrValArgThrSerLeuPro</td> <td>28</td> </td>	21 <td>ArgTyrValArgThrSerLeuPro</td> <td>28</td>	ArgTyrValArgThrSerLeuPro	28
b <td>61 <td>CGTTATGTCGGCAACTCTATCTC-ACGGTAAGAAGTTGAACACACAATTTATTTGTTTA</td> <td>119</td> </td>	61 <td>CGTTATGTCGGCAACTCTATCTC-ACGGTAAGAAGTTGAACACACAATTTATTTGTTTA</td> <td>119</td>	CGTTATGTCGGCAACTCTATCTC-ACGGTAAGAAGTTGAACACACAATTTATTTGTTTA	119
Y <td>28</td> <td></td> <td>28</td>	28		28
b <td>120 <td>ATGTTTTCAITGGTAACTAGAGTCTTAAAACTTAGCCTAGACGATACACAGCATCTT</td> <td>179</td> </td>	120 <td>ATGTTTTCAITGGTAACTAGAGTCTTAAAACTTAGCCTAGACGATACACAGCATCTT</td> <td>179</td>	ATGTTTTCAITGGTAACTAGAGTCTTAAAACTTAGCCTAGACGATACACAGCATCTT	179
Y <td>28</td> <td></td> <td>28</td>	28		28
b <td>180 <td>GAATCTAGATTCAATATTTATACAGAAATATTTATTTTAAATACGATATAGTTCCAG</td> <td>239</td> </td>	180 <td>GAATCTAGATTCAATATTTATACAGAAATATTTATTTTAAATACGATATAGTTCCAG</td> <td>239</td>	GAATCTAGATTCAATATTTATACAGAAATATTTATTTTAAATACGATATAGTTCCAG	239
Y <td>28</td> <td></td> <td>28</td>	28		28
b <td>240 <td>ATTTTAAATTTGGTACATAAGAAAGATACTAGATTCTAACGAAATTAACCACTTGCA</td> <td>299</td> </td>	240 <td>ATTTTAAATTTGGTACATAAGAAAGATACTAGATTCTAACGAAATTAACCACTTGCA</td> <td>299</td>	ATTTTAAATTTGGTACATAAGAAAGATACTAGATTCTAACGAAATTAACCACTTGCA	299
Y <td>28</td> <td></td> <td>28</td>	28		28
b <td>300 <td>CTGAAAGATCCGAGCATAAAGTGTGTTACTATATAAAGAGTATTTCTTTTTTAATCTTA</td> <td>359</td> </td>	300 <td>CTGAAAGATCCGAGCATAAAGTGTGTTACTATATAAAGAGTATTTCTTTTTTAATCTTA</td> <td>359</td>	CTGAAAGATCCGAGCATAAAGTGTGTTACTATATAAAGAGTATTTCTTTTTTAATCTTA	359
Y <td>29</td> <td></td> <td>40</td>	29		40
b <td>360 <td>AGCTAAATATCAATTTTTCATCAGATTCGAAATACCTTAGAATCTGATCCCTTAGGNA</td> <td>419</td> </td>	360 <td>AGCTAAATATCAATTTTTCATCAGATTCGAAATACCTTAGAATCTGATCCCTTAGGNA</td> <td>419</td>	AGCTAAATATCAATTTTTCATCAGATTCGAAATACCTTAGAATCTGATCCCTTAGGNA	419
Y <td>41 <td>AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu</td> <td>60</td> </td>	41 <td>AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu</td> <td>60</td>	AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu	60
b <td>420 <td>GCAGCATCAAAATCATCAACGACGACTCAAGTTTGACGGTAAACCCGAGGCTAAACCTG</td> <td>479</td> </td>	420 <td>GCAGCATCAAAATCATCAACGACGACTCAAGTTTGACGGTAAACCCGAGGCTAAACCTG</td> <td>479</td>	GCAGCATCAAAATCATCAACGACGACTCAAGTTTGACGGTAAACCCGAGGCTAAACCTG	479
Y <td>61 <td>AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle</td> <td>80</td> </td>	61 <td>AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle</td> <td>80</td>	AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle	80
b <td>480 <td>GCCTCCTTTGTGACCACTGGATGGAGCAGAAATGTGACRAGCTCATGATGAAATCCATC</td> <td>539</td> </td>	480 <td>GCCTCCTTTGTGACCACTGGATGGAGCAGAAATGTGACRAGCTCATGATGAAATCCATC</td> <td>539</td>	GCCTCCTTTGTGACCACTGGATGGAGCAGAAATGTGACRAGCTCATGATGAAATCCATC	539
Y <td>81 <td>AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys</td> <td>100</td> </td>	81 <td>AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys</td> <td>100</td>	AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys	100
b <td>540 <td>ACAAGAAACAACGTTGAGATGGACCAATACCTTACCACCGACCTTCAGAAATCGATGC</td> <td>599</td> </td>	540 <td>ACAAGAAACAACGTTGAGATGGACCAATACCTTACCACCGACCTTCAGAAATCGATGC</td> <td>599</td>	ACAAGAAACAACGTTGAGATGGACCAATACCTTACCACCGACCTTCAGAAATCGATGC	599
Y <td>101 <td>ValAsnMetIleAlaHisIleuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly</td> <td>120</td> </td>	101 <td>ValAsnMetIleAlaHisIleuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly</td> <td>120</td>	ValAsnMetIleAlaHisIleuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly	120
b <td>600 <td>GTTAATCATGTTGCGGCTCTTTCAACGGCCCTTTAGGTGACGGTGAAGCCGCCATTTGGT</td> <td>659</td> </td>	600 <td>GTTAATCATGTTGCGGCTCTTTCAACGGCCCTTTAGGTGACGGTGAAGCCGCCATTTGGT</td> <td>659</td>	GTTAATCATGTTGCGGCTCTTTCAACGGCCCTTTAGGTGACGGTGAAGCCGCCATTTGGT	659
Y <td>121 <td>ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys</td> <td>140</td> </td>	121 <td>ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys</td> <td>140</td>	ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys	140
b <td>660 <td>GTTGGCAGGTTGGGGTTCATCGGAGGAGTGTGGCCGGACTGGCCCTTTAAGAGACAG</td> <td>719</td> </td>	660 <td>GTTGGCAGGTTGGGGTTCATCGGAGGAGTGTGGCCGGACTGGCCCTTTAAGAGACAG</td> <td>719</td>	GTTGGCAGGTTGGGGTTCATCGGAGGAGTGTGGCCGGACTGGCCCTTTAAGAGACAG	719
Y <td>141 <td>TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly</td> <td>160</td> </td>	141 <td>TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly</td> <td>160</td>	TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly	160
b <td>720 <td>TGGCAGAACAGGTAAGCCCTTAGGGCTGCCCTTATGATAGACCTTAATTTGTAACCGGA</td> <td>779</td> </td>	720 <td>TGGCAGAACAGGTAAGCCCTTAGGGCTGCCCTTATGATAGACCTTAATTTGTAACCGGA</td> <td>779</td>	TGGCAGAACAGGTAAGCCCTTAGGGCTGCCCTTATGATAGACCTTAATTTGTAACCGGA	779
Y <td>161 <td>AlaAsn</td> <td>162</td> </td>	161 <td>AlaAsn</td> <td>162</td>	AlaAsn	162
b <td>780 <td>GCCAATATTCAGGTAAACCAAAAACAAAATTTGATTAATTTTAAACCGGTTTAGGCTAT</td> <td>839</td> </td>	780 <td>GCCAATATTCAGGTAAACCAAAAACAAAATTTGATTAATTTTAAACCGGTTTAGGCTAT</td> <td>839</td>	GCCAATATTCAGGTAAACCAAAAACAAAATTTGATTAATTTTAAACCGGTTTAGGCTAT	839
Y <td>163</td> <td></td> <td>173</td>	163		173
b <td>840 <td>GTTTACATGACTCAATTTCCGGTTCAATACAGAGTTTGGTTGGAGAAATTTGCAAGGTAT</td> <td>899</td> </td>	840 <td>GTTTACATGACTCAATTTCCGGTTCAATACAGAGTTTGGTTGGAGAAATTTGCAAGGTAT</td> <td>899</td>	GTTTACATGACTCAATTTCCGGTTCAATACAGAGTTTGGTTGGAGAAATTTGCAAGGTAT	899
Y <td>174 <td>PheGluValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGln</td> <td>193</td> </td>	174 <td>PheGluValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGln</td> <td>193</td>	PheGluValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGln	193
b <td>900 <td>TTTGAAGTGGAGCTTAAAGRATGAGCTGAGASAGGATATACGTTGATGGACCCCTGAC</td> <td>959</td> </td>	900 <td>TTTGAAGTGGAGCTTAAAGRATGAGCTGAGASAGGATATACGTTGATGGACCCCTGAC</td> <td>959</td>	TTTGAAGTGGAGCTTAAAGRATGAGCTGAGASAGGATATACGTTGATGGACCCCTGAC	959

Alignment Scores:
 red. No.: 2493
 1834.50
 Matches: 395
 Conservative: 44
 Mismatches: 52
 Similarity: 70.15%
 Gaps: 6

180 uValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAs 200
 837 GGTAAACTAAGTGAAGACTACTATGTTATGGATCCAGTAAAGCTGAGAGTGGTGA 896
 200 pGluAsnThrIleCysValAlaAspIleuGlySerThrLeuAsnGlyGluPheGluAs 220
 897 TGAGAAATACCACTGTTGGCAATCTTAGATCCACACTTACTGGAGAGTTTGAGGA 956
 220 pValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrPr 240
 957 CGTAAAGCAATTAAGCAATCTTCTAGTGGAGAAAACGACAGACAGATGGAAATCC 1016
 240 oIleHisValAspAlaAserGlyPheIleAlaProPheLeuTyrProGluLeuG1 260
 1017 TATTGATGTTGAGCAGCCAGTGGAGATCAATGCTCTCTCCCTCAACCCGATTTGA 1076
 260 uTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLe 280
 1077 ALGGGACTTTAGCTTCCATGGTGAAGAGTATTAACGTCAGTGTCAAGATATGGACT 1136
 280 uValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeu 300
 1137 TGTGATGAGGAGTGGTGGTGGTGGTGGAGAAACAAAGATGATTTGCCAGAGAACT 1196
 300 uIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerIy 320
 1197 TGTCTTCCACATCAACTACTTGGGAGCTGATCAACCCACTTTCACCTCAACTTCTCAA 1256
 320 s ----- 320
 1257 AGGTTGTPAAATAAAACTGGCTTTATCCAAATCAATCCATCAATCAATTTCCTTTAA 1316
 321 -----GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuI1 333
 1317 AAACCTCAATGTTTTCTTTTTCAGGGTCGAGCCAAATCAATGCTCAGTACTACTACT 1376
 333 eArgLeuGlyHis----- 337
 1377 CCGACTAGGCTT-TGAGGTACTTGTTCCTTATFCGCAATACAGTTTCACTCT 1435
 338 -----GluGlyTyrArgAsnValMetGluAs 346
 1436 TGCCTTAATCTAATGATCTTTTGGAACTGAAAGGGATACAGAACATTAATGGRAAA 1495
 346 nCysArgGluAsnMetIleValIleuArgGluGlyLeuGluLysThrGluArgPheAsnI1 366
 1496 CTGCATGGATAACGCCAAGGAGGCTAAGAGAGCAATAGAGATGACAGGGAAGTTCACAT 1555
 366 eValSerLysAspGluGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHi 386
 1556 TGTGTCCTCAAGATATTGGCGTCCACTAGTGGCATTCCTCTCAAAAGACAGTAGCAAGCA 1615
 386 sThrGluPheGluIleSerAspMetLeuArgTyrGlyTyrIleValProAlaTyrTh 406
 1616 CAGGCTGTTGAGATCGCAGACTTTTGAGAAAATTCGGGTGGATCATACCGGCTTACAC 1675
 406 rMetProProAsnAlaGlnHisIleThrValLeuArgValValIleArgGluAspPheSe 426
 1675 TATGCTGCAGATGCACAGACATCTGCTGCTCAGAGTTGTGATAAGAGAAAGACTTTAG 1735
 426 rArgThrLeuAlaGluArgLeuValIleAspIleGluLysValMetArgGluLeuAspG1 446
 1736 CCGAGCCCTTGCAGATAGACTCATACATATCAATTCAGGTCGCTGAAAAGAGATTGAAG 1795
 446 uLeuProSerArgValIleHisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSe 466
 1796 GCTTCTAGCAGGATTCACATCTTCTGCTGG-----GCTGCACCGGTTAGTGG 1843
 466 rAspAsnLeuMetValThrValLysLysSerAspIleAspLysGlnArgAspIleIleTh 486
 1844 TGATGATGAAGAAATTAAGTGAAGACTGCAAGATGCTCCTTGGAG---GATATCACTAA 1900

Qy 486 xGlyTyrLysPheValAlaAspArgLys 496
 Db 1901 GTATTGGAACCCCTTGTGGAACACACAAGAGA 1931
 RESULT 15
 AAC44889
 ID AAC44889 standard; DNA; 1607 BP.
 XX AAC44889;
 XX 18-OCT-2000 (first entry)
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 44511.
 XX Hybridisation assay; genetic mapping; gene expression control;
 XX protein identification; signal transduction pathway;
 XX metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 FN EP1033405-A2.
 XX 06-SEP-2000.
 PF 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121925.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.


```

18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
28-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144684.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
05-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
06-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.

```

```

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment Scores:
Pred. No.: 1.85e-153
Score: 1570.00
Percent Similarity: 69.64%
Best Local Similarity: 61.27%
Query Match: 50.04%
DB: 21
Length: 1607
Matches: 337
Conservative: 46
Mismatch: 51
Indels: 117
Gaps: 10

```

```

US-10-006-852-2 (1-502) x AAC444889 (1-1607)

```

```

Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer

```

1 ATGGTTTGGACAAAACCGCAACG---AATGATGAATCTGTCTGCACCACATGTTCCGGATCT 57
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 58 CGCTATGTTCCGACTACACTTCCCAAGTATGAGATTGGTGAATTCGATACCAGAAC 117
 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 118 GCTGGCTATCAGATCAATAAAGATGAGTGTCTGAT--- 156
 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 157 -----GAGTGTGACAAACTCATCATGGACTCTATC 186
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln----- 97
 187 RAAT---AACTAGCTTGATATGATGATGATACCTGTCCACAACTGAGCTCAGTGTGATGAA 243
 97 ----- 97
 244 AGTTCATGTGGCGTAGTGGTAATACTCCRAAATGCATTTCTCCACCAGGGGTATAAT 303
 98 -----Asn 98
 304 CGTAATTAACACAACTTCTTCTCCRAACAAACCCTAGTAGTCCGCTTCTCTTAAC 363
 99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAla 118
 364 CGATGTGTAACCAATATAGCTCGACTGTTCAATCGCCACTCGAGGAATCTGAGACGGCG 423
 119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
 424 GTGGGAGTAGGGACAGTGGTGTCTTCAGAAAGCCATCATGTGTAGCCGGATGGCCCTCAA 483
 139 ArgLysTyrGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleVal 158
 484 AGAAATGGCAGAAACAACGCAGAGCTGAGGGTAACCCCTATGACAAACCCCAATTTGT 543
 159 ThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeu 178
 544 CACCGGATCCCAATCAAGTTTCTGGGAGAAATTCGCTCGGTACTTCAGGTGGAG--- 600
 179 LysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnAlaValAspMet 198
 601 CAAGGAGTAAACCTAAGTGAAGGTTACTACGCTGATGGATGATCCAGACAAAGCAAG 660
 199 ValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPhe 218
 661 GTAGACGAGAAACAATCTGTGTGCA---ATATTTGGGATCCACACTCAAC----- 708
 219 GluAspValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAsp 238
 708 ----- 708
 239 ThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGlu 258
 709 -----GCAGTGGGTCTAGCTCCGTTTATCTATCTGAA 744
 259 LeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyr 278
 745 TTAGAATGGACTTTAGGCTTCTTTGGTTAAGAGCATCAACGTGAGTGGTCCACAAGTAT 804
 279 GlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGlu 298
 805 GGACTAGTCTATCTGCTGGTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 864
 299 GluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPhe 318
 865 GAGCTTATCTTTTCATATTAATATATCTTGGTGTGATCAACCCACTTTCCTACTCAATTC 924
 319 SerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGln 338
 925 TCCAA--GGATCCGAGCCAAATATTGCTCAATACTACAGCTCATTTCCGCTCTTGGAA---- 978

Qy 338 uGlyTyrArgAsnValMetGluAsnCysArgGlu-AsnMetIleValLeuArgGluGlyL 358
 979 -----TTCGAGAATTCATAGAAACAACATGGTGGTTCCTCAAAAGAGGTA 1022
 Qy 358 euGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaLap 378
 1023 TAGAGAAAAACAGAGCCGTTTCAACATAGTCTCAAAAGACCACAAAGGAGTGCACGTCGTCGCT 1082
 Qy 378 he-SerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArg 397
 1083 TCITCTCTCAAGACCACATAGTTCCACACAGAGTTCAGATCTCTGAGATGCTACGTCGT 1142
 Qy 398 TyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeu 417
 1143 TTRGGCTGGATCGTCCAGCTTACACTATGCTTCCTTCCAAGATATCTAAGAAGATGGGAAT 1322
 Qy 418 ArgValValIleArgGluAspPheSerArgThrLeuAlaGlu-ArgLeuValIleAspI 437
 1203 CGTGTGTCTCATAGGAAGATTTCTCAAGAAACTCCGGAAGAGACTTGTGCTGATAT 1262
 Qy 437 eGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLe 457
 1263 TTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAGATATCTAAGAAGATGGGAAT 1322
 Qy 457 uGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAs 477
 1323 A-----GAAAGGATCGCGGAAAAATGTAAGGAGAGAGAA 1355
 Qy 477 pIleAspLysGln-----ArgAspIleIleThrGlyTrpLysLysPheValAlaAspAr 495
 1356 GATGGAGAGGAGATTTCTGATGGAAAGTATTTGTTGATGGAGGAAGTTTGTGAGGAGAG 1415
 Qy 495 gLysLysThrSerGlyIleCys 502
 1416 GAAGAAGATGAATGGTGTGTGC 1437

Search completed: October 22, 2003, 15:31:10
 Job time : 406 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - nucleic search, using frame_plus_p2n model
run on: October 22, 2003, 15:09:39 ; Search time 4557 Seconds
(without alignments)
4506.612 Million cell updates/sec

US-10-006-852-2
erfect score: 2615
equence: 1 MWLSHAVBSYSDVSHSTFAS.....DIITGWKRFVADRKKTGIC 502

coreing table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

searched: 2889711 segs, 20454813386 residues
total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

command line parameters:
MODEL=frame+pnz.model -DEV=xlp
Q=/cgn2_1/USPTO.spool_p/US10006852/runat_21102003_164103_2453/app_query.fasta_1.647
DB=GenEmbl -QFWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0
UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cgi -LIST=45
DOCALIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFWT=pc0 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US10006852@cgn_1_4958@runat_21102003_164103_2453 -NCPU=6 -ICPU=3
MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

database :

- GenEmbl:*
- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_pr:*
- 7: gb_ro:*
- 8: gb_sy:*
- 9: gb_un:*
- 10: gb_vl:*
- 11: em_ba:*
- 12: em_fun:*
- 13: em_hum:*
- 14: em_in:*
- 15: em_mu:*
- 16: em_or:*
- 17: em_ov:*
- 18: em_ph:*
- 19: em_pi:*
- 20: em_ro:*
- 21: em_sts:*
- 22: em_un:*
- 23: em_ba:*
- 24: em_fun:*
- 25: em_hum:*
- 26: em_in:*
- 27: em_mu:*
- 28: em_or:*
- 29: em_ov:*
- 30: em_ph:*
- 31: em_pi:*
- 32: em_ro:*
- 33: em_sts:*
- 34: em_un:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2615	100.0	1509	8	ATU10034	U10034 Arabidopsis
2	2607	99.7	1509	6	AX507115	AX507115 Sequence
3	2607	99.7	1509	6	AX651552	AX651552 Sequence
4	2607	99.7	1813	8	BT001047	BT001047 Arabidops
5	2607	99.7	1813	8	AY094464	AY094464 Arabidops
6	2250	86.0	1785	8	PETGADX	L16797 Petunia hyb
7	2250	86.0	1785	8	PETGLUDECA	L16797 Petunia hyb
8	2240	85.7	1705	8	AF020425	AF020425 Nicotiana
9	2233	85.4	1745	8	AF506366	AF506366 Nicotiana
10	2232	85.4	1929	8	NTU54774	US4774 Nicotiana t
11	2228	85.2	1672	8	AF352732	AF352732 Nicotiana
12	2200	84.1	1479	6	AX654603	AX654603 Sequence
13	2200	84.1	1502	8	AY187941	AY187941 Oryza sat
14	2199	84.1	1771	8	AF020424	AF020424 Nicotiana
15	2186.5	83.6	1776	8	AF353615	AF353615 Nicotiana
16	2135.5	81.7	1482	6	AX412801	AX412801 Sequence
17	2135.5	81.7	1482	6	AX506242	AX506242 Sequence
18	2135.5	81.7	1482	6	AX651853	AX651853 Sequence
19	2135.5	81.7	1482	8	AY124860	AY124860 Arabidops
20	2135.5	81.7	1792	8	AF361836	AF361836 Arabidops
21	2130	81.5	1485	8	AY124873	AY124873 Arabidops
22	2130	81.5	1676	8	ATU46665	U46665 Arabidopsis
23	2130	81.5	1681	8	ATU49937	U49937 Arabidopsis
24	2130	81.5	1695	8	AF428372	AF428372 Arabidops
25	2126	81.3	1685	8	AF428294	AF428294 Arabidops
26	2102	80.4	1650	8	AY081259	AY081259 Arabidops
27	2080	79.5	1712	8	AKL18125	AKL18125 Arabidops
28	2065.5	79.0	75188	8	AB005238	AB005238 Arabidops
29	2056.5	78.6	1779	8	AB056060	AB056060 Oryza sat
30	2014	77.0	1783	6	A39964	A39964 Sequence 8
31	2014	77.0	1783	6	AR070945	AR070945 Sequence
32	2007	76.7	1455	6	LEGDL	X80840 L.esulentu
33	1965	75.1	1605	8	AX653206	AX653206 Sequence
34	1965	75.1	1605	8	AY084689	AY084689 Arabidops
35	1945	74.4	136098	8	AC137696	AC137696 Genomic S
36	1936.5	74.1	76976	8	AC006532	AC006532 Arabidops
37	1894	72.4	147227	8	OSJN00013	OSJN00013
38	1854	70.9	1467	6	AX653024	AX653024 Sequence
39	1844	70.5	1781	8	AB056061	AB056061 Oryza sat
40	1836	70.2	1180	6	BD017613	BD017613 Method fo
41	1782	68.1	27194	8	AB026646	AB026646 Arabidops
42	1699	65.0	147227	8	OSJN00013	OSJN00013
43	1696	64.9	3983	8	AB056063	AB056063 Oryza sat
44	1630	62.3	5160	6	AX461514	AX461514 Sequence
45	1630	62.3	81490	8	FLP19	AC009513 Sequence

ALIGNMENTS

ATU10034 1509 bp mRNA linear PLN 30-MAR-1996
 Arabidopsis thaliana Columbia glutamate decarboxylase (GAD) mRNA, complete cds.
 U10034
 U10034.1 GI:497978
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 1509)
 Arazi,T., Baum,G., Snedden,W.A., Shelp,B.J. and Fromm,H.
 Molecular and biochemical analysis of calmodulin interactions with the calmodulin-binding domain of plant glutamate decarboxylase
 Plant Physiol. 108 (2), 551-561 (1995)
 95334488
 7610159
 2 (bases 1 to 1509)
 Fromm,H.
 Direct Submission
 Submitted (25-MAY-1994) Hillel Fromm, Department of Plant Genetics, Weizmann Institute of Science, Rehovot 76100, Israel
 Location/Qualifiers
 1..1509
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strains="Columbia"
 /db_xref="taxon:3702"
 /clone="358377"
 1..1509
 /gene="GAD"
 1..1509
 /gene="GAD"
 /EC_number="4.1.1.15"
 /note="GAD1 isoform; similar to the GAD2 isoform of Arabidopsis thaliana, GenBank Accession Number U49937"
 /codon_start=1
 /product="glutamate decarboxylase"
 /protein_id="AAA93132.1"
 /db_xref="GI:497978"
 /translation="MVLGHVSESDVSHSTFASRYVYRSLPRFKMPENSIPKBAAYQ IINDELDMPRLNLSFVITWMEPCDKLMSINKNYDMDDEYVPTTELQKRCVN MIAHFNAPLBEAETAVGVGVGSESLMLAGLAFKRWKQKRAEKPVDKPNIVG ANVQVWEKFAFYFEVLEKVKLSEGYVYMDPQQADVMDENTICVADILGSTLNGF EDVLLNDLLVKKKETEWDTPIHVDAASGGFIAPFLYPELEWDFRPLVKSINVSFH KYGLVYAGIGWVWRKEDLPEELIFHNYLGDQPTFLNFKSGSOVIAOYYLIR LQHGVRVNMENCRNMLVLRGKLEKTERFNI VSKDEGVLVAFSLKDSCHTEFELS DMLRRYGNVPAVYMPNAQSHITVLRVIREDFRTEAERLVLDIEKVMRELDLPSR VIKHLSLGGKESNSDMLMVTVRKSDIDKQDDITGWKKEVADRKRKTSYGIC"
 misc_feature 1417..1506
 /gene="GAD"
 /note="calmodulin-binding domain"
 CASE COUNT 433 a 321 C 385 g 370 t
 ALIGNMENT
 Alignment Scores:
 Score: 2615.00 Length: 1509
 E-value: 100.00 Matches: 502
 Identity: 100.00 Conservative: 0
 Positives: 100.00
 Negatives: 0
 Gaps: 0
 S-S-10-006-852-2 (1-502) x ATU10034 (1-1509)
 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 |||||
 1 ARGGGCTCTCCACCGCGTATCGGAGTGGAGCTCCGTCCTCCACTCCACATTCGCATCA 60
 |||||
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 |||||

Db 61 CGTTACGTCGCTACTTCACCTTCTAGTTCAAGATCCCGGAAACTCGATTCTTAAGGAA 120
 Qy 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 |||||
 Db 121 CGGGGTATCAGATCATCAACGACGAGCTGATCTTACGGGGAATCCACGGTTGAACCTA 180
 Qy 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 |||||
 Db 181 GCCTCTTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATATGCTCTCCATC 240
 Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 |||||
 Db 241 AACAGAACCTATGTTGACATGGACGATACCCCGTCCACCACCGAATTCAGAACCGATGT 300
 Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThraAlaValGly 120
 |||||
 Db 301 GTGAACATGATTCACATCTATTCAATGCACCGTTAGAGAGCGGGAGACCGCGCTCGGA 360
 Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 |||||
 Db 361 GTAGGAACCGTTGGATCATCGGAGGCCATAAATGTTGGCCGTTTGGCCCTTCAAGCGTAAA 420
 Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 |||||
 Db 421 TGGCAGAACAGCGCAAGCTGAAAGCCAAACCCGTCGATAAACCCCAACATTTGTCACCGGA 480
 Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 |||||
 Db 481 GCCAATGTTCAAGTGTGTTGGAGAAATTCGTAGTACTTTGAGGTGAACCTTAAGGAA 540
 Qy 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnIleAlaValAspMetValAsp 200
 |||||
 Db 541 GTGAAATGAGTGAAGGATATGATGAGCCCTCAACCAAGCTGTGATATGGTGTAT 600
 Qy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 |||||
 Db 601 GAGAACCACTTGTGTTGGCGACATCTTGGTTCACCTCTTAAATGGAGAAATTCGAAAT 660
 Qy 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 |||||
 Db 661 GTTAAACTTGAACCGATCTCTTGGTCGAAAAGAAACAAAGAAACCGGATGGGATACACCA 720
 Qy 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 |||||
 Db 721 ATCCAGTGGATCGGCAAGTGGAGGATTCATTCGACCGTTTTTGTATCCGGAATGGAA 780
 Qy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 |||||
 Db 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGATCAATGTGAGTGGTCAACAAGTATGGACTT 840
 Qy 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeuLeu 300
 |||||
 Db 841 GTGTACGCGGGATTTGGTGGTGAATCTGGAGAAAACAAAGAGGATTTGCCCTGAGGAATC 900
 Qy 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlyProThrPheThrLeuAsnPheSerLys 320
 |||||
 Db 901 ATCTTCCATATCAATATCTTGGTGTGACCAACCCACCCCTTTACTCTCAATTTCTCCAAA 960
 Qy 321 GlySerSerGlnValIleAlaGlnTyrTrpGluLeuIleArgLeuGlyHisGluGlyTyr 340
 |||||
 Db 961 GGTTCAGCTCAGTCTTGTCTCAATACTACCAACTTATCCGATTTGGCCACGAGGGTTAC 1020
 Qy 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 |||||
 Db 1021 AGAAATGTGATGGAAATTCAGAGAAATATGATCGTCTAAGGGAAGGACTTGAGAAG 1080
 Qy 361 ThrGluArgPheAsnIleValSerLysAspGluClyValProLeuValAlaPheSerLeu 380
 |||||
 Db 1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCACCTTGTCTTCTCTCTTTG 1140
 Qy 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTyr 400
 |||||
 Db 1141 AAGATAGCAGCTGTCACTGATTCGAANATCTCCGACATGCTTCCAGGTAATGATGG 1200

401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1201 ATAGTCCGGCCTACAAATGCCCTCCAAATGCACAAACACATCACTGTTCCTCGTGGTT 1260
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1261 ATCAGAGAAGATTTCTCGAAGCACTCGCTCGAGACTTGTGATCGATATAGAAAAGTG 1320
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1321 ATGCGTGAAGCTCGATGACCTCCCTCGAGAGTATCAAAAATATCACTTGGCAAGAG 1380
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1381 AAGAGTGAATCTAACAGCGATAAATGATGTCACCGTGAAGAAAGCGCATATCGACAAG 1440
 481 GlnArgAspIleIleThrClyTrpLysLysPheValAlaAspArgLysThrSerGly 500
 1441 CAGAGAGATATCATCCTGGCTGGGAAGATTTGCGCCGACAGGAAGACGAGTGGT 1500
 501 IleCys 502
 1501 ATCTGC 1506

SULT 2
 507115
 CUS
 AX507115 1509 bp DNA linear PAT 27-SEP-2002
 Sequence 1810 from Patent WO0216655.
 AX507115
 AX507115.1 GI:23388352

Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1
 Harper,J.F., Krebs,J., Wang,X. and Zhu,T.
 Stress-regulated genes of plants, transgenic plants containing
 same, and methods of use
 Patent: WO 0216655-A 1810 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG
 (CH)

FEATURES
 source Location/Qualifiers
 1..1509
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 432 a 322 c 385 g 370 t

BASE COUNT
 432 a 322 c 385 g 370 t

ALIGNMENT
 Alignment Scores:
 Length: 1509
 Score: 2607.00 Matches: 501
 Percent Similarity: 99.80% Conservative: 0
 Local Similarity: 99.80% Mismatches: 1
 Identity Match: 99.69% Indels: 0
 Gaps: 6

1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 1 ATGGTCTCTCCACCGCGTATCGGAGTCGGAGCTCCGTCCCTCCATCCACATTCGCATCA 60
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 61 CGTTAGCTCCGATCTCACTTCTCTAGTTCAAGATGCGGAAAATCGATTCTTAAGGAA 120
 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgGluAsnLeu 60
 121 GCGGGGTATCAGATCATCAACGACGAGCTGTGCTTGGCGGAAATCCACGGTTGAACITTA 180

Qy 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCTGCTCCATC 240
 Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 Db 241 AACAAAGAACATGTTGACATGGACGAGTACCCCGTCCACCAGAACTTCAGAACCCATGT 300
 Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluAlaGluThrAlaValGly 120
 Db 301 GTGAAACATGATGTCACATCTATTCATATGCACCGTTAGAGAGCGGAGACCCCGTGG 360
 Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 361 GTAGGAACCGTTCGATCATCGAGGCGCATAAATGTTGGCCGTTTGGCCCTCAAGCGTAA 420
 Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 Db 421 TGGCAGAACAAAGCGCAAGCTGAAGGCAACCCCGTGCATAAACCCCAACATFTGFCACCG 480
 Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 481 GCCAATGTTCAAGTGTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGAA 540
 Qy 181 ValLysLeuSerGluGlyTyrValMetAspProGlnAlaValAspMetValAsp 200
 Db 541 GTGAATTTGAGTGAAGGATACTATGTGATGGACCCCTCAACAGCTTGTGATATGGTTGAT 600
 Qy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 Db 601 GAGAACACCAATTTGTTGGGCCATTTCTTGGTTCACCTCTTAATGGAAATTCGAAGAT 660
 Qy 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 Db 661 GTTAAACTTTCGACGATCTCTTGGTCGAAAGAACAAAGAACCCGATGGGATACACCA 720
 Qy 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 Db 721 ATCCACGTGATCGGACGAGTGGAGGATTCATTCACCCGTTTTTGTATCCGGAATGGAA 780
 Qy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 Db 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGATCAATGTGAGTGGTCAAAAGTATGACACT 840
 Qy 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 Db 841 GTGTACCGCAGGATTTGTTGGTGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAATC 900
 Qy 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 Db 901 ATCTTCCATATCAATATCTTGGTGTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
 Qy 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 Db 961 GGTTCAAAGTCAAGTCAATGTCTCAATCACTCACTATCCGATTTGGCCACGAGGGTTAC 1020
 Qy 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 Db 1021 AGAAATGTGATGAGAAATTCAGAGAAATATGATCGTCCCTAAAGGAAAGGACTTGAAG 1080
 Qy 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 Db 1081 ACAGAAAGTTCAACATCGTCTCAAAGGACGAGGAGTGCACCTTGTTCGCTTCTCCTTG 1140
 Qy 381 LysAspSerSerCysHisThrGluPheGluIleSerAsnMetLeuArgArgTyrGlyTrp 400
 Db 1141 AAAGATAGCAGCTGCACACTGAGTTCGAAATCTCCGACATGCTTCGAGGATGAGTGG 1200
 Qy 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 Db 1201 ATAGTCCGGCCTACAAATGCCCTCCAAATGCACAAACACATCACTGTTCCTCGTGGTT 1260

421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1261 ATCAGAGAAGATTTCTCGAGAACAACCTCGCTGAGACATTTGTGATCGATATAGAGAAAGTG 1320
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1321 ATGCGTGGCTCGATGAGTTCCTTCGAGAGTCAATCAAAATATCACTTGGACAAAG 1380
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1381 AAGAGTGAATCTAACAGCGATAAATGTATGTCTACCGGTGAAGAGCGGATATCGACAAG 1440
 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
 1441 CAGAGAGATATCATCTGCTGGTGGGAAGAAGTTTGTGGCCGACAGGAAGACGAGTGGT 1500

501 IleCys 502
 1501 ATCTGC 1506
 AX651552 1509 bp DNA linear PAT 22-MAR-2003
 Sequence 360 from Patent WO03000898.
 AX651552
 AX651552.1 GI:29154370

Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, S., Tac, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
 Plant genes involved in defense against pathogens
 Patent: WO 03000898-A 360 03-JAN-2003;
 Syngenta Participations AG (CH)

FEATURES
 source 1..1509
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 432 a 322 c 385 g 370 t

Alignment Scores:
 aligned. No.: 1509
 core: 2607.00 Matches: 501
 percent Similarity: 99.80% Conservative: 0
 Local Similarity: 99.80% Mismatches: 1
 Best Local Match: 99.69% Indels: 0
 Gaps: 6

1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 1 ATGGTCTCTCCACCGCTATCGGAGTGGACGCTCTCGTCCACTCCACATTCGCATCA 60
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 61 CGTTACGTCGACTTCACTTCTTCTAGTTTCAAGTCCGGAACCTCGAATTCCTAAGGAA 120
 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 121 GCGGCGTATCAGATCATCAACGACGAGCTGATCGTTCGCGGAATCCACGGTGTGACTTA 180
 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 181 GCCTCCTTTGTGACGACATGGATGGAGCCTGTGATGATAAACTCATATGTCCTCCATC 240
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100

Db 241 AACAAAGAAATATGTTGATCGAGAGTACCCCGTACCACCGAACTTCAGAACCCGATGT 300
 Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 Db 301 GTGAACATGATTCACATCTATTCAATGCACCGTTAGAAAGGCGGAGACCGCCGTCGGA 360
 Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 361 GTAGAAACCGTGGATCATCGGAGCCATAATGTTGGCCGTTTGGCCCTTCAAGCGTAAA 420
 Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 Db 421 TGGCAGAACAGCCCAAGCTGAGGCGAAACCCCGTGCATAAACCCCAACATTTGCACCGGA 480
 Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 Db 481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGTAGGTACTTTGAGGTTGAACCTAAGGAA 540
 Qy 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 Db 541 GTGAAATTTGATGAAAGGATATATGATGAGACCCCTCAACAAGCTGTTGATATGGTTGAT 600
 Qy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 Db 601 GAGAACCAATTTGTGTTGGCCCAATTTCTGGTTCCACTCTTAAATGGAGAAATTCGAAGAT 660
 Qy 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTyrAspThrPro 240
 Db 661 GTTAAACTTGTGAACGATCTCTTGGTCGAAAAGAAACAAGAAACCCGGATGGATACACCA 720
 Qy 241 IleHisValAspAlaIleSerGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 Db 721 ATCCACGTGATCGGCAAGTGGAGGATTCATTCACCCGTTTTTGTATCCGGAATTCGAA 780
 Qy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysLysTyrGlyLeu 280
 Db 781 TGGACATTTAGACTTCCCTTGGTGGAGAGATCAATGTGAGTGTCAAGATATGGACTT 840
 Qy 281 ValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 Db 841 GTGTACCGAGGATTTGGTGGTGTCTGGAGAAAACAAGAGGATTTGCTCGAGAACTC 900
 Qy 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 Db 901 ATCTTCCATATCAATTAFTTGGTGTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
 Qy 321 GlySerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 Db 961 GGTTCAGTCAAGTCAATGCTCAATCACTCAACTTATCCGATTCGCGCCACGAGGTTAC 1020
 Qy 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlyLeuGluLys 360
 1021 AGAAATGTGATGGAGAAATTCAGAGAGAAATATGATCGTCTTAAGGGAAGGACTTGAGAAG 1080
 Qy 361 ThrGluArgPheAsnIleValSerLysAspGlyValProLeuValAlaPheSerLeu 380
 Db 1081 ACAGAAAGTTTCAACATCGTCTCAAGAGGACGAGGAGTGCCTTTCGCTTCTCCTTTG 1140
 Qy 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgAspTyrGlyTyr 400
 Db 1141 AAGATAGCAGCTGTCACTGAGTTCGAAATCTCCGACATGCTTCGCGAGTATGGATGG 1200
 Qy 401 IleValProLafThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 Db 1201 ATAGTCCGGCCCTACAAATGCCCTCCAAATGCAACAACATCACTGTTCTTCGTGCTGT 1260
 Qy 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 Db 1261 ATCAGAGAAGATTTCTCGAAGAACCTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
 Qy 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460

1321 ATGCGTGAAGCTCGATGAGCTTCCCTTCGAGAGTGATTCACAAANTATCACTTGGACAAGAG 1380
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1381 AAGAGTGAATCTAACACGGATAAATCTGATGCTCAGCGTGAAGAGGCGATATCGCAAG 1440
 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
 1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTGCGCGACAGGAAGACGAGTGGT 1500
 501 IleCys 502
 1501 ATCTGC 1506

BT001047 1509 bp mRNA linear PLN 18-OCT-2002
 Arabidopsis thaliana At5g17330/MKP11_18 mRNA, complete cds.
 BT001047
 BT001047.1 GI:24111354
 FLI CDNA.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1509)
 Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,
 Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
 Palm, C.J., Quach, H.L., Sakurai, T., Sacou, M., Seki, M., Southwick, A.,
 Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K.,
 Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
 Ecker, J.R.

Arabidopsis ORF clones
 Unpublished
 2 (bases 1 to 1509)
 Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,
 Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
 Palm, C.J., Quach, H.L., Sakurai, T., Sacou, M., Seki, M., Southwick, A.,
 Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K.,
 Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
 Ecker, J.R.

Direct Submission
 Submitted (18-OCT-2002) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEZ (SSP) Consortium members constructed and
 sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk, R.,
 Chen, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Chan, M.M.,
 Chang, B., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S.,
 Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M.,
 Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A.,
 and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.
 Location/Qualifiers
 1. .1509
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"

/db_xref="taxon:3702"
 /chromosome="15"
 /clone="U21967"
 /notes="this clone is in pUNI 51
 ecotype: Columbia"
 1. .1509
 /notes="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
 /codon_start=1
 /product="At5g17330/MKP11_18"
 /protein_id="AAM46801.1"
 /db_xref="GI:24111355"
 /translations="MVLSHAVSESDVSHVTFASRVVRLSLPRFKMPENSIKPEAAAYQ
 IIADELMDGNENLNLASVFTTWMEPECDKLIINSKKNYVDWDEYFVTVTELQNRQVNI
 MIHFLNAPLEAEATAVGTVGSSSEATMLAGLAFKRWKQKKAEGKPKVRENIVTG
 ANVQVCWEKFAFYFEVLEKLVKSEGVYVMDPQQAVMDVMDENTICVAALIGSLNGEFL
 EDKVLNDLLVKNKTKGWDTPIHDAASGGFTAPFLYPELWDFRFLPLVKSINVSCH
 KYGLVAGIGMVIWRNKEDLPEELIPIHNYLGDQPTFTLNFSGSSQVTAQYVQLIR
 LGHEGRVNVNENRENMIVLREGLEKTERFNIVSKGEPVLFVAFSLKSSCHTFEFLFR
 DMLRRYGVIPAYTPMPNAQHITVLRVWIDRFSRLAERLVIDIEKVRMRELDLFSR
 VIHKISLQBEKSESNDNLMVTYKKSIDIKQRDIITGWKKFVADRKKTSGIC"
 BASE COUNT 432 a 322 c 385 g 370 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.63e-211 Length: 1509
 Score: 2607.00 Matches: 501
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.69% Indels: 0
 DB: 8 Gaps: 0

US-10-006-852-2 (1-502) x BT001047 (1-1509)
 Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 Db 1 ATGGTGTCTCCACCGCCGATCGAGTCCGAGCGTCTCCCTCCACTCCACATTCGCAATCA 60

Qy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 Db 61 CGTTACCTCCGACTTCACTTCCCTAGGTTCAAGATGCCGAAAACCTCGATTCCTAAGAA 120

Qy 41 AlaAlaTyrGlnIleIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 Db 121 GCGCGGTATCAGATCAATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

Qy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 Db 181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATPAACTCATCATGTCTCCATC 240

Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 Db 241 AACAAAGAACTATGTTGACATGGACGAGTACCCCGCTCACCCAGAACTTCAGAACCCGATGT 300
 Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
 Db 301 GTGAACATGATTTGCACATCTATTCAATGCACCGTATTAGAGGGCGGAGACCGCCGTCGGA 360

Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 Db 361 GTAGAACCCGTTGGATCATCGAGGCCATTAATGTTGGCCGGTTTGGCCITTCAGCGTAAA 420
 Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 Db 421 TGGCAGAACAAAGCCAAAGCTGAAGGCAACCCGCTCGATAAACCCCAACATTTGTCCACCGGA 480
 Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLysGlu 180
 Db 481 GCCAAATGTTCAAGTGTGTTGGGAGAAATTCGTAGGTACTTTGAGGTTGAACCTTAAGGAA 540
 Qy 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 Db 541 GTGAATTTGAGTGAAGGATACTATGTGATGAGACCCCTCAACAGCTGTTGATGTTGAT 600

CDS

CDS

RESULT 4

LOCUS 0001047

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 |||||
 601 GAGAACACCATTTGGTGGCCATCTTGGTCCACTTAAATGGAGAAATCCAGAGT 660
 |||||
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 |||||
 661 GTPAAACTCTTGAACGATCTTGGTGGAAAGAACAAAGAAACCGATGGATACACCA 720
 |||||
 241 IleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 |||||
 721 ATCCACGTGGATGGCCAAAGTGGAGATTCATTGACCGCTTTTGTATCCGGAATTGAA 780
 |||||
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 |||||
 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCAAGATATGGACTT 840
 |||||
 281 ValTyrAlaGlyIleGlyTrpValIleTrpAspAsnLysGluAspLeuProGluLeuLeu 300
 |||||
 841 GTGTACCGCAGGATGGTGGTGGTGGTGGAGAACAAAGAGGATTTCCCTGAGGAATC 900
 |||||
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 |||||
 901 ATCTTCCATATCAATATCTTGGTGGTGGACCAACCCACTTACTCTCAATTTCTCCAAA 960
 |||||
 321 GlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 |||||
 961 GGTTCAAAGTCAAGTCAATCTCAACTACTACCAACTATCCGATTCGGCCACGAGGTTAC 1020
 |||||
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuLys 360
 |||||
 1021 AGAATGTGATGGAAATTCAGAGAGATATGATCGTCTTAAAGGAGGACTTGGAGAG 1080
 |||||
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 |||||
 1081 ACAGAAAGGTTCAACATCTCTCAAAAGGACGAGGAGTCCACTTGTCTCTCTCTCTG 1140
 |||||
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
 |||||
 1141 AAAGATGACGAGCTGTCCACTGAGTTCGAAATCTCCGACATCTTCCGAGGATGGATGG 1200
 |||||
 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 |||||
 1201 ATAGTCCCGCCCTACCAATGCTCCAAATGCACCAACACATCACTGTTCTTCTGTTGT 1260
 |||||
 421 IleArgLysAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 |||||
 1261 ATCAGAGAAAGATTTCTCGAGAACACTCGCTGAGACACTTGTGATCGATATAGAGAAAGT 1320
 |||||
 441 MetArgLysLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 |||||
 1321 ATCGTGGAGCTCGATGAGCTTCTTCGAGAGTGAATTCAAAATATCACTTGGACAGAG 1380
 |||||
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 |||||
 1381 AAGAGTCAATTAACAGCGATAACTTGTGAGTCAACCGTGAAGAGAGGATATCGACAA 1440
 |||||
 481 GlnArgAspIleIleThrGlyTrpLysPheValAlaAspArgLysLysThrSerGly 500
 |||||
 1441 CAGAGAGATATCATCACTCGCTGGAGAAAGTGTGTCGCCGACAGGAGGACGAGTGGT 1500
 |||||
 501 IleCys 502
 |||||
 1501 ATCTGC 1506

RESULT 5
 Y094464
 DEFINITION 1813 bp mRNA linear PLN 05-MAY-2002
 ACCESSION AY094464
 ERSION AY094464.1 GI:20453186
 KEYWORDS FLI CDNA,
 Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (bases 1 to 1813)
 AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 TITLE Arabidopsis cDNA clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1813)
 AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 TITLE Direct Submission
 JOURNAL Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Kamiya, A., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

FEATURES
 source 1..1813
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAF109-48-C10 (R21967)"
 /note="ecotype: Columbia"
 74..1582
 /note="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
 /codon_start=1
 /product="AT5G17330/MKP11_18"
 /protein_id="AA019834.1"
 /db_xref="GI:20453187"
 /translations="MVLSHAVSSDYSVHSTFASRYVRTSLPRFKMSPENISPKAAAYO IINDELMDGNPELNLASFVTTWMEPECDKLMSSINKVVDVMDVPTVELQNRCTV MIARLNPALREAEATVAVGTGVSSEAIIMLAGLAFKRNQKKAEGKQVDFXNVTV ANVQWCKEFPARYFVELKEVLEGGYVMDPQAVDMDVDEMTICVAALIGSTLNGEF EDVKLLNDLLEKKNKQETGWDTPHVDAAAGGTFAPLELWDFELPKVFNYSVGH KYGLVAGIGWIVRNKQDPELHINYLADQPTFTLNFSSKSSQSVIAQYVQILR LGHEGYNTMENCENNI VLREKLEKTRFNI VSKDEGVPVAVSLKDSSTCHTEPFI DMLRRVGVIPVATPEPNAQIHIVLVRVIEDEPSPRLAERLVIDIEKRWRELDLPSR VIHKSILGQKSESNDLMTVTKKDIQRDIITGKWKFFVADRKTKTSGIC"

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.
 Location/Qualifiers

BASE COUNT 504 a 384 c 417 g 508 t
 ORIGIN
 Alignment Scores:

red. No.: 2,04e-211 Length: 1813
 core: 2607.00 Matches: 501
 percent Similarity: 99.80% Conservative: 0
 est Local Similarity: 99.80% Mismatches: 1
 query Match: 99.69% Indels: 0
 B: 8 Gaps: 0

S-10-006-852-2 (1-502) x AY094464 (1-1813)

1 MetValIleuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 74 ATGGTCTCTCCACCGCCGATCGGAGTCGGAGTCCTCCGTCACCTCCACATTCGGATCA 133
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 134 CGTACGTCGACTTCCATCTCAGGTTCAAGATCCCGAAAACTCGATTCCTAAGGAA 193
 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 194 CGGCGTATCAGATCAACACGACGAGTCGATCGTTGACGGGAATCCACGGTTGAACCTA 253
 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 254 GCCTCCCTTGTGACGACATGATGGAGCCGTGATGTGATAACTCATCATGTCCTCCATC 313
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
 314 AACAGAACTATGTGATGATGACGAGTACCCCGTCACCCAGAACTTCAGAAACCGATGT 373
 101 ValAsnMetIleAlaHisIlePheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 374 GTGAACATGATTCACATCTATTTCATGTCACCCGTTAGAGAGCGGAGACCCTCCGTCGA 433
 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 434 GTAGAAACCGTTCGATCATCGGAGGCCATAATGTTGGCCGTTTGGCCCTTCAAGGGTAA 493
 141 TrpGlnAsnLysArgLysAlaGluLysProValAspLysProAsnIleValThrGly 160
 494 TGCAGAACAAAGCCGAAAGCTGAAGCCAAACCCCGTCGATAAACCCCAACATGTCACCCG 553
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 554 GCCAATGTTCAAGTGTGGGAGAAATTCCTGAGTACTTGGAGTTGAACTTAAGGAR 613
 181 ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
 614 GTGAAATGTAGTGAAGGATACTATGTGATGGACCCCTCAACAAGCTTTCATATGTTGAT 673
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrIleAsnGlyGluPheGluAsp 220
 674 GAGAACACCAATTTGTTGGCCGCAATTCCTGGTTCACCTTAAATGGAAATTCGAAGAT 733
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 734 GTTAAACTCTTGAACGATCTCTTGTTCGAAAGAACAAAGAAACCGATGGGATACACCA 793
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 794 ATCCACGTGATCGGCAAGTGGAGATTCATGTCACCGTTTTTGTATCCGGAATTCGAA 853
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 854 TGGACATTTAGACTTCCCTTGGTGAAGATCAATGTGATGGTTCACAAAGTATGGACTT 913
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 914 GTGTACCGAGGATGTTGGTGTATCTGGAGAAACAAAGAGGATTTGCTGAGGAACTC 973
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 974 ATCTTCCATATCAATATCTTGGTGTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 1033

321 GlySerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 1034 GGTTCAGTCAAGTCAATGCTCAATCACTCACTCACTCACTCACTCACTCACTCACTCACT 1093
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 1094 AGAAATGTGATGGAGAAATTCAGAGAGATATGATCGTCTTAAGGAAGGACTTGAAG 1153
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 1154 ACAGAAAGTTCACATCGTCTCAAGACAGAGGAGTCCACTTTCGCTTTCTCCTTG 1213
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTyr 400
 1214 AAAGATAGCAGCTGCTCACACTGAGTTCGAAATCTCCGACATGCTTCGAGTATGGATGG 1273
 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1274 ATAGTCCCGCCCTACACAATGCTCCAAATGTCACACACATCACTGTTCTTCGTGGT 1333
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1334 ATCAGAGAGATTTCTCGAGACACCTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1393
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1394 ATGCGTGGAGCTCGATGAGTTCCTTCGAGAGTGAATTCACAAAATATCACTTGGACAAG 1453
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1454 AAGAGTGAATCTAACACCGATACTTGTGATGCTCACGGTGAAGAGAGCGATATCGACAAG 1513
 481 GlnArgAspIleIleThrGlyTyrLysPheValAlaAspArgLysLysThrSerGly 500
 1514 CAGAGAGATATCACTCACTCGCTGGAGAGTTCCTCCGACAGAGAAAGACGAGTGGT 1573
 501 IleCys 502
 1574 ATCTGC 1579

RESULT 6
 PETGDX
 LOCUS Petunia hybrida glutamate decarboxylase (gad) mRNA, complete cds.
 DEFINITION
 ACCESSION L16797
 VERSION L16797.1 GI:294111
 KEYWORDS glutamate decarboxylase.
 SOURCE Petunia x hybrida
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterididae; Lamiales; Solanales; Solanaceae; Petunia.

REFERENCE 1 (bases 1 to 1785)
 AUTHORS Baum,G., Chen,Y., Arazi,T., Takatsuji,H. and Fromm,H.
 TITLE A plant glutamate decarboxylase containing a calmodulin binding domain. Cloning, sequence, and functional analysis
 JOURNAL J. Biol. Chem. 268 (26), 19610-19617 (1993)
 MEDLINE 93374956
 PUBMED 8366104
 COMMENT Original source text: Petunia hybrida flower corolla cdna to mRNA.
 FEATURES Location/Qualifiers
 source 1..1785
 /organism="Petunia x hybrida"
 /mol_type="mRNA"
 /db_xref="taxon:4102"
 /tissue_type="corolla"
 /dev_stage="flower"
 gene 1..1785
 /gene="gad"
 CDS 72..1574
 /EC_number="4.1.1.15"
 /codon_start=1

```

/product="glutamate decarboxylase"
/protein_id="AAA33709.1"
/db_xref="GI:294112"
/translation="WVLSKTVSQSDVSIHSTFASRYVRSIPLRFKMPDNPISPKAAAYQ
IINDELMDGNPRLMLASFTVYVMEPECDKLMDSINKNYVDMDEYFVPTTELQNRVW
MAHLFNAPLEDEGTAVGVTGSSSIALAGLAFKRWQNRKAKGKCDKPKNIYVG
ANVQVWEKFAFYFEVLEKVEVSEYVMPPEKAVENVDENTICVAAILGSLTNGBF
EVKRLNDLLVKNKTKMDTPIHVDAAAGGPIAPPIYPELWDFRPLVKSINVSQH
KYGLVYAGLGVWVWNRKNDLPDLFIHYNLIGADOPTFLNFSKGSQSOVIAQYYOLIR
LGEYGVKNVMEQENASVLRSEGLKTRFNILISKEIGVPLVASLKDNRQHNFEIS
ETIIRRFGLVPAITPPNACHLIVURVIREDFSKTLAERLVRDLKVLVHELDLILPAR
VNRKLAVEQAAANGSEVHKHTDSEVQLEMLTAWKKEVEBKKTNRVC"

```

```

ASE COUNT 550 a 298 c 401 g 536 t
RIGIN
Alignment Scores:
red. No.: 3, 5e-181 Length: 1785
core: 2250.00 Matches: 427
Percent Similarity: 92.64% Conservative: 39
Best Local Similarity: 84.89% Mismatches: 33
Very Match: 86.04% Indels: 4
Gaps: 2
3-10-006-852-2 (1-502) x PETGADX (1-1785)

```

/	1	MetValLeuSerHisAlaValSerGluSerAspValSerValHiserThrPheAlaSer	20
/	72	ATGGTTCATCAAAGACAGTCCAGAGCGATGTGTCATTCACCTCCACGTTTCTCT	131
/	21	ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu	40
/	132	CGATATGTTCGAACCTCTCTCCAGGTTTAAATAAGCCAGATAAATTCGATACCAAGAA	191
/	41	AlaAlaTyrGlnIleleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu	60
/	192	GCAGCATATCAGATCATAATGATGACATGATGTAGATGGAACCCCAAGCGTGAACITG	251
/	61	AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle	80
/	252	GCITCTTTGTTCAACATGGATGGAACACAGAGTGTGATAAGTTGATGATGACCTATT	311
/	81	AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys	100
/	312	AACAAGAACATGATGATGGAAGAAATFCCCTGTTACCACTGAGCTTCAGAAATCGATG	371
/	101	ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly	120
/	372	GTAACATGATAGCTCATTGTTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA	431
/	121	ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys	140
/	432	GTGGAACTGTTGGATCCTCTGAAGCCATTATCTGCTGGATTAGCTTCAAGAGAAA	491
/	141	TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly	160
/	492	TGGCAGAACAAAATGAAAGCCCAAGCCCTGTGCAAGCCCAACATGTTACTGCT	551
/	161	AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu	180
/	552	GCRAATCTCCAGTGTCTCGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGAA	611
/	181	ValLysLeuSerGluGlyTyrValMetAspProGlnGlnAlaValAspMetValAsp	200
/	612	GTAAGCTTAGTGAAGACTATGATGACCCCTGGAAGACTGTGAGATGGTGGAT	671
/	201	GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp	220
/	672	GAAACACCAATTTGTGTAGCTGCTATCTTAGTTCCACCTCAATGGAGAAATTTGAAGAC	731
/	221	ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro	240
/	732	GTTAAGCGCTTGAATGATCTCTTGGTCGAGAGAAACAAAGAAACCGGGTGGACATCCA	791

Qy	241	IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu	260
Db	792	ATTCAATGGATGCACCAAGTGGTGGATTTATGGACCGTTCATTTACCAGAGCTTGAG	851
Qy	261	TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu	280
Db	852	TGGGACTTTAGATTGGCATTAGTGAAGACATAAATGTAAGTGGTGCACAAAATATGGTCTT	911
Qy	281	ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu	300
Db	912	GTCATGCTGGTATTGGTGGGTGGTTCGGAGGAAACAAGATGATTTGCTGATGAACCTT	971
Qy	301	IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys	320
Db	972	ATCTTCCACATTAATTAATCTTGGTGTGATCAACCTACTTTCACTCTCAACTTTCTAAA	1031
Qy	321	GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr	340
Db	1032	GGTTCAGCCCAAGTAAATGCTCAATTAATACCAACTTATTCGCTGGGTATGAGGGTTAC	1091
Qy	341	ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlyLys	360
Db	1092	AAGAACTGTGATGAGAAATGCTCAAGAAAATGTCATCGGTACTAAGAGAAGGGCTAGA	1151
Qy	361	ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu	380
Db	1152	ACAGGAAGATTCACCAATAATCTCCAAGAAAATGGAGTACCTTTAGTAGCATTTCTCTT	1211
Qy	381	LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTyr	400
Db	1212	AAGACCAACAGGCAACACACAGATTCGAGATTTCTGAAACTTTAAGAGAGATTTGGTGG	1271
Qy	401	IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal	420
Db	1272	ATTGTTCTCATATACTATGTCACCAAAACGCAACACACATTCAGTTCTCAGAGTTGTG	1331
Qy	421	IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal	440
Db	1332	ATCAGAAAGATTTCTCCCTGTCAGAGACGACTGTTAAGACATATCGAAAAGTTC	1391
Qy	441	MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu	460
Db	1392	CTTCATGAACTTGACACACTCCCTGACAGTCTCAATGCTAAGCTGCTGGCCGAGGAG	1451
Qy	461	LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys	480
Db	1452	CAGCCGCTCGAATGCGACCGAGGTCAT-----AAGAAAACAGATAGCGAAGTG	1502
Qy	481	GlnArgAspIleIleThrGlyTyrLysLysPheValAlaAsp---ArgLysLysThrSer	499
Db	1503	CAGTTGGAGATGATGCTGCAATGGAGAGCTTTGTTCAAGAAAAGAAAGAAAGACTAAT	1562
Qy	500	GlyIleCys 502	
Db	1563	CGAGTTTGT 1571	
RESULT 7			
PETGLUDECA			
LOCUS	Petunia hybrida glutamate decarboxylase (gad) mRNA	linear	PLN 26-JUL-1993
DEFINITION			
ACCESSION	L16977		
VERSION	L16977.1	GI:309679	
KEYWORDS	calmodulin binding protein; glutamate decarboxylase.		
SOURCE	Petunia x hybrida		
ORGANISM	Petunia x hybrida		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	Asteridae; Lamiales; Solanales; Solanaceae; Petunia.		
JOURNAL	1 (bases 1 to 1785)		
	Baum, S., Chen, Y., Arazi, T., Takatsuki, H. and Fromm, H.		
	A plant glutamate decarboxylase containing a calmodulin-binding		
	domain: cloning sequence and functional analysis		
	J. Biol. Chem. (1993) In press		

COMMENT Original source text: Petunia hybrida flower corolla cDNA to mRNA.

FEATURES source location/Qualifiers
/organism="Petunia x hybrida"
/db_xref="taxon:4102"
/tissue_type="corolla"
/dev_stage="flower"
/gene="gad"
CDS
/gene="gad"
/EC number="4.1.1.15"
/cofon_start=1
/product="glutamate decarboxylase"
/protein_id="AAA33710.1"
/db_xref="gi:309680"
/translations="MVLSTKVSQDYSIHSTFASRYRTSLPRFKVDPNSIPKAAAYQ
IINDELMIDGNRLNLSAFVTTWMEPECDKLMDSINKNYVDMDEYFVPTTELQNRVNV
MIAHLFNAPLEDEETAVGVGTGSEALMLAGLAKRKNWKNKQKPCDKPNTVTG
ANVQWKEFARYPEVELKEVLSKGYVMDPEKAVEMVDENI CVAA ILGSLTNGBF
EDVKRLDLLVEKNKSTGWDTPRHVDAASGGBTAPEIYSELEWDRPLPLVKSINVSCH
KYGLVYAGICGVWRNKDLDLDELIFHINFLGADQPTFLNFSKSSQVIAGYQIIR
LGEGYKVMNENCOENASVREGLEKTRFNIIISKEIGVPLVAFSLKDNQRHNEFEIS
ETLRFEGWI VPATYPPNPAOHITLVLRVIREDFSRFLAERLVRDLKVLHELDLTPAR
VNAKLVAEBOAAANGSEVHKHTDSEVQLEMITAWKKFVEEKKKTRVC"
ASE COUNT 550 a 298 c 401 g 536 t
RIGIN

Alignment Scores:
red. No.: 3.5e-181 Length: 1785
core: 2250.00 Matches: 427
percent Similarity: 92.64% Conservatative: 39
est Local Similarity: 84.89% Mismatches: 33
very Match: 86.04% Indels: 4
B: 8 Gaps: 2

S-10-006-852-2 (1-502) x PETGLUDECA (1-1785)
1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
72 ATGGTTCTATCAAGACAGTGTGCGAGAGCGAGTGTCCATTCCACCCAGTTTGTCTT 131
21 ArgTyrValArgThrSerLeuProArgPheIysMetProGluAsnSerIleProLysGlu 40
132 CGATATGTCGAACTTCTTCCAGGTTTAAAATGCGAGATAATTCGATACAAAAGAA 191
41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
192 CCAGCATATCAGATCAATATGATGAACTGATGTAGATGGAACCCAGGCTGAACTTG 251
61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
252 GCTTCTTTTGTACAACTGATGAAACCCAGGCTGATGATGATGATGATGATGATGAT 311
81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
312 AACAGAACTAATGTTGATGGAATGAAATTCCTGTTACCCTGAGCTTCCAGATCGATG 371
101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
372 GTAAACATGATAGTCTATTTGTTAAATGCACCCTTGAAGATGGAGAACTGCGAGTTGGA 431
121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
432 GTTGGAACTGTTGGATCTCTGAAAGCCATATGCTTCTGGATGATGATGATGATGATG 491
141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
492 TGGCAGACAAAATGAAAGCCCAAGGCAACCTGTGACAAAGCCCAACTTGTACTGGT 551
161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180

552 GCAAAATGTCAGGTTGCTCGGAGAAATTTGCAAGGTAATTTTGAAGTGGAGCTAAAGGAA 611
181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
612 GTAAGGTTTATGAAGGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 671
201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
672 GAAACACACATTTGTAGCTGCTATCTTAGTTCCACCTCAATGGAGAAATTTGAGAC 731
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
732 GTTAAAGCGCTTGAATGATCTTGGTCGAGAAACAAAGAAACCGGTTGGACACTCCA 791
241 IleHisValAspAlaAlaSerGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
792 ATTCATGTCGATGAGCAGAGTGTGGATTTATGACCTTCAATTTACCAGAGCTTGG 851
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
852 TGGACATTTAGATTTGCCATTTAGTGAAGAGCATTAATGTAAGTGGTCAAAATATGCTCT 911
281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
912 GTCATGCTGGTATTTGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 971
301 IlePheHisIleAsnTyrLeuGlyValAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
972 ATCTTCCACATTAATATCTTGGTGTGATCAACTTACTTCACTCTCAACTTCTCTAAA 1031
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
1032 GGTTCAGCAAGTAAATTTGCTCAATATTAACCACTTATTCCGTTGGTGGTGGTGGTGG 1091
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGluLys 360
1092 AAGATGTCATGTCGAAATTTGTCAGAAATGTCATCGTACTAGAGAGAGGCTTGAARAG 1151
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
1152 ACAGGAAGATTCACATAAATCTCCAAAGAAATTTGGAGTACCTTTAGTAGCAATTTCTCT 1211
381 LysAspSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
1212 AAAGACACAGGCAACACAGAGTTCGAGATTTTGAACCTTTAAGAGATTTGGTTGG 1271
401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
1272 ATTTGTCCTGATATCTATGCCCAACCCACACACACACATTAACAGTTCTCAGAGTTGTG 1331
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
1332 ATCAGAAAGATTTCTCCCGTACCTTGCAGACGACTGGTAAAGACACATPCGAAAAGTC 1391
441 MetArgLysLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1392 CTTTCATGACTTGCACACTCCCTGACAGTTCATGCTAAGTCCGCTGCTGGCCGAGGAG 1451
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1452 CAGGCGGCTGCGAATGGCAGCGAGTGCAT-----AAGAAAACAGATAGCGAAGTG 1502
481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp--ArgLysLysThrSer 499
1503 CAGTTGGAGATGATCTGATGGAAGAGTTTGTGAAGAAAAGAGAAAGAAAGAACTAAT 1562
500 GlyIleCys 502
1563 CGAGTTTGT 1571
RESULT 8
AF020425 1705 bp mRNA linear PLN 24-JUN-1998
LOCUS AF020425

DEFINITION Nicotiana tabacum glutamate decarboxylase isozyme 1 (NtGAD1) mRNA, complete cds.

CCRESSION AF020425

ERSON AF020425.1 GI:3252855

YWORDS Nicotiana tabacum (common tobacco)

ORGANISM Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 1705) Yun, S.J. and Oh, S.H.

AUTHORS Cloning and characterization of a tobacco cDNA encoding calcium/calmodulin-dependent glutamate decarboxylase

TITLE Unpublished

JOURNAL 2 (bases 1 to 1705) Yun, S.J.

REFERENCE Direct Submission

AUTHORS Submitted (21-AUG-1997) Faculty of Biological Resources Science, Chonbuk National University, 664-14 1-ga Tokjin-dong, Chonju, Chonbuk 561-756, South Korea

JOURNAL Location/Qualifiers

TITLE 1 .1705

FEATURES source

1..1705 /organism="Nicotiana tabacum"

mol_type="mRNA"

/db_xref="taxon:4097"

1..1705 /gene="NtGAD1"

71..1561 /gene="NtGAD1"

/functions="calmodulin binding protein"

/note="calcium-calmodulin-dependent enzyme"

/codon_start=1

/product="glutamate decarboxylase isozyme 1"

/protein_id="AAC24195.1"

/db_xref="GI:3252856"

/translations="MVLSTAESDYSIIHSTFASRYVVRSLPRFKMPENSIPKEAAYQ IINDELMDDGNPRINLASFVITWBEFECNKLMMDSINKNYVDMDEYPTVEIQNRQVNI MIARFLAFLDGETAVGVTGSSSEALMLAGLAFKRWKQKAGKPCDKPNIIVTG ANVQCVKRFARFEVLEKELSDGYYVMDPEKAVEMVDENIICVAALGSLNNGE EDVKRLNDLLIEKNKGTWDPFHVDAAGSGFPAPELPELJEDWDFLPLVKGINVSGH KGLVYAGIGIWAWRNKEDLPBLIFHINVLGADQPTFLNFKGSSQVIAQYVOLLIR LGREGYKWNVEQENARVLRLEKESGRNIIKELIYPLVAFSLKDNKSOHNEFEIS ETLARFQWIIIPATYMPNAGQVLRVLRVREDFSRILABELVIDEKVLHEDLILPAR VNAKLAVAERANGSVEVHKTKDREYQLETTIYAKKFKVADKXKXKNGVC"

ASE COUNT 523 a 301 c 385 g 496 t

IGIN

Alignment Scores:

red. No.: 2.32e-180 Length: 1705

Score: 2240.00 Matches: 428

Percent Similarity: 92.05% Conservative: 35

Best Local Similarity: 85.09% Mismatches: 32

Very Match: 85.66% Indels: 8

3: 8 Gaps: 3

Qy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100

Dd 311 AACAGAAATACCGTGGAGTGAATACCCCTGGTAAACCACTGAACCTTCAGAAATCGAATG 370

Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120

Dd 371 GTAACATGATAGTCACTTTGTTAAAGCACCACCTGGAGATGGAGAGACTCAGTTGCA 430

Qy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140

Dd 431 GTTGAACCTGTTGGATCCTCGAGGTATATGCTTCTGGATTAGCTTTCAGAGAAAA 490

Qy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160

Dd 491 TGGCAAAATTAATGAAAGCCCAAGCCCTGTGACAAAGCCCAATATTTGTCACCTGGT 550

Qy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180

Dd 551 GCCAAATGCCAGGTGTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610

Qy 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200

Dd 611 GTAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670

Qy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220

Dd 671 GAGAACACAATTTGTGTAGCTGCTATCTTTGGTTCACACTCAATGGTGAATTTGAAAGAT 730

Qy 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrClyTrpAspThrPro 240

Dd 731 GTTAAGCGCTTGAATGACCTCTTGTGATGAGAAACAAAGAAACCGGGTGGGACACTCCA 790

Qy 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260

Dd 791 ATTCATGTGGATGCAGCAAGTGTGGATTTATTTGCACCACTTCTTTATCCAGAGCTTGA 850

Qy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280

Dd 851 TGGGACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910

Qy 281 ValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeuLeu 300

Dd 911 GTTTATGCTGGTATGTTGGCCATTTGGAGAAATAAGGAGACTTACTGACGAACTT 970

Qy 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320

Dd 971 ATCTTCCACATTAATTTATCTTGGTGTGATCAACCTTCTTCACTTCTCACTTCTTCTAAA 1030

Qy 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340

Dd 1031 GGTTCCTAGCCAAAGTAAATGCTCAATTAATCAACCACTTATTCGCTGGGTTTGGAGGTTAC 1090

Qy 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360

Dd 1091 AAGAAATGTTATGGAGAAATTTGCAAGAAATGCAAGAGGCTACTAAGAGAGAGACTTGA 1150

Qy 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380

Dd 1151 AGTGGAGATTCACATTAATCCAAAGAAATTTGGATTCCTTCCATTAGTGTCTTCTCTCT 1210

Qy 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400

Dd 1211 AAGACACACAGTCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1270

Qy 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420

Dd 1271 ATTATTCCTGCATATATATGCCCAAAATGCTCAACATGCTCACAGTCTTCAGAGTTGTC 1330

Qy 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440

Dd 1331 ATTAGAGAAAGATTTCTCCCTTACTCCTCCGAGGACTGTTAATATGACATTTGAAAGTTC 1390

```

441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1391 CTCCACGAGTAGACACACTTCGGCGAGGGTCAACGCTAAGCTAGCCGGT----- 1441
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1442 ---GCCGAGCGGMAATGCCAGCGCGTGCATF-----AAGAAACAGATAGAGAAGTG 1489
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1490 CAGCTTGAGATTACTACTCATCGAAGAAATTTGTTGCTGATTAAGAGAGAACTAAC 1549
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
500 GlyIleCys 502
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1550 GGAGTTTGT 1558

AF506366      1745 bp mRNA linear FIN 02-MAY-2003
Nicotiana tabacum glutamate decarboxylase mRNA, complete cds.
AF506366
AF506366.1 GI:21327028
Nicotiana tabacum (common tobacco)
Nicotiana tabacum
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 1745)
McLean, M.D., Yevtushenko, D., Deschene, A., Van Cauwenbergh, O. R.,
Makhmoudova, A., Potter, J.W., Bown, A.W. and Shelp, B.J.
Overexpression of glutamate decarboxylase in transgenic tobacco
confers resistance to the northern root-knot nematode
Mol. Breed. 11, 277-285 (2003)
2 (bases 1 to 1745)
McLean, M.D., Yevtushenko, D. and Shelp, B.J.
Direct Submission
Submitted (25-APR-2002) Plant Agriculture, University of Guelph,
Guelph, ON N1G 2W1, Canada
FEATURES
    Location/Qualifiers
        1..1745
            /organism="Nicotiana tabacum"
            /mol_type="mRNA"
            /cultivar="Samsun NN"
            /db_xref="taxon:4097"
            3'..1521
                /note="GAD20.1"
                /codon_start=1
                /product="glutamate decarboxylase"
                /protein_id="AAM48129.1"
                /db_xref="GI:21327029"
                /translation="MVLSTKASESDVSIHSTFASRYVFTSLPRFKMPENSIKPEAAVQ
                IINDELMDGNPRMLASVFTTMMPECCNLMDSINKNYVDMDEPVTTLQNRQVN
                MIAHLNAPLGGGTAVGVTGSSSEAILMGLAFKRWKQKRAQGFPPKPNIVTG
                ANVQWKEFAYFVELKVELSDGYVMDPEKAVEMVDENI CVAAIILGSLNGE
                EOVKRLNDLLEKNETQWDTPIHVDAAAGSFFIAPPLYLEWDRPLPLVKSINVS
                KYGLYAGLGMWLNENKDLPELIFHNYLGADQPTFLNFKSSSVIAQYQLLR
                LGFEGKVMQNCNVRVREGIEKSGRFNLSKEIGVPLVAFSLKDNSQHNFEFIS
                ETLRFQWTPAYTPPNAQHVTYLVREDFSLRBLRVLDIEKVLHELDLTPAR
                VNAKLVAEANGSGVHKKTRDREVLQLEITAWKKFVADKKKKTVGC"
    ASE COUNT      555 a 290 c 389 g 511 t
    ALIGN

Alignment Scores:
    -red. No.:      9.37e-180      Length:      1745
    score:          2233.00      Matches:     427
    percent Similarity: 91.85%      Conservative: 35
    best Local Similarity: 84.89%      Mismatches:  33
    Query Match:     85.39%      Indels:      8
    3:                8          Gaps:        3

3-10-006-852-2 (1-502) x AF506366 (1-1745)

```

```

Oy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 31 ATGTTCTGTCCAAAGACAGCGTCCGAAAGTGCAGTCTCCATCCACTCCACTTCGTTCC 90
Oy 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 91 CGATATGTTCCAACTCTCTCCAGGTTTAAAGTCCAGAGAATTCATACCAAGGAA 150
Oy 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 151 GCAGCATATCAGATTATAATGATGAGCTTATGTTAGATGGAATCCAGGCTAAATTA 210
Oy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
Db 211 GCATCTTTCGTACAACATGATGGAGCCAGAAATGTAATACGTTAATGATGGATTCCA 270
Oy 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
Db 271 ACAAGAACTACGTTGACATGGATGAATACCCGTAAACCACTGAGCTTCAGAAATCGAT 330
Oy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
Db 331 GTAATAATGATAGCTCACTTGTGTTTAAATGCACCACTTGGAGATGGAGACTGCAGTT 390
Oy 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
Db 391 GTTGGAACTCTGGATCTCTGAAGCTATTATGCTTGTGGATTAGCTTTAAAGAAAA 450
Oy 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
Db 451 TGGCAAAATTAATGAAAGCCCAAGCCACTTGTATAGCCCAATATCTGCCCGGT 510
Oy 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
Db 511 GCTAATGTCAGGCTGTGTTGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 570
Oy 181 ValLysLeuSerGluGlyTyrValMetAspProGlnAlaValAspMetValAsp 200
Db 571 GTAATAATGATGATGGATACTATGATGGAGCCCTGAGAAAGCTGTGGAATTTGGT 630
Oy 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
Db 631 GAGAAATACCAATTTGTTGCTGCTACTCTTAGGTTCAACTCAATGTTGAAATTTGAGAT 690
Oy 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
Db 691 GTTAAAGCGTTTGAATGACCTTTTCAITGAGAGAAACAAGAAACAAACCGGGTGGACA 750
Oy 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
Db 751 ATTCATGTGATGAGCAAGTGTGGATTATTCACCACTTCCTTTATCCAGAGCTTGA 810
Oy 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
Db 811 TGGGACTTTAGATTGCCATTGGTGAAGAGTATTAAATGAGTGTGTCACAAAATATGCTCT 870
Oy 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
Db 871 GTCTATGCTGGTATTGGTTGGCCCAITTTGGAGGATAAGAGAGACTTGCCTGATGA 930
Oy 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnGlnPheSerLys 320
Db 931 ATTTTCCCACTCAATTTACCTTGTGCTGATCAACCTTCTACTCTCAACTTCTCTAAA 990
Oy 321 GlySerSerGlnValIleAlaGlnTyrTrpGlnLeuIleArgLeuGlyHisGluGlyTyr 340
Db 991 GGTTCAGCCCAAGTAATTTGCTCAATATTACCACTATTTCGCTGGTTTGGAGGGTTAC 1050
Oy 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
Db 1051 AAGAATGTTATGAGAATTTGTCAAGAAAAATGCAAGGGTATTAAAGAGAGAAATTTGAAA 1110
Oy 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380

```

```

...
1111 AGTGGAGATTCAACATAATCCAAAGAAATGGAGTTCCCTTAGTAGCATTTCCTT 1170
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
1171 AAAGCAACAGTCAACACAAATGAGTTCGAAATTTCTGAACTTTAGAGATTTGGATGG 1230
401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
1231 ATTGTTCTTCATATACTATGTCACCAAAATGCTCAACATGTTACAGTTCACAGATTCTC 1290
421 IleArgGluAspPheSerArgThrIleuAlaGluArgLeuValIleAspIleGluLysVal 440
1291 ATTAGAGAAGATTTCTCCGACACATAGCGGAGCCTGGTAAATAGACATTAAGAAAGTC 1350
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1351 CTCACGACTAGACACACTTCGGGGAGGTCACGCTAGCTAGCCGCTG----- 1401
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1402 ---GCCGAGCGAATGCGACGCGCTGCAT-----AAGAAAACAGATAGAGAAGTG 1449
481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
1450 CAGCTAGAGATTAAGTACTGCTGATGGAAGAAATTTGTTGCTGATAGAAAGAAAGACTAA 1509
500 GlyIleCys 502
1510 GGAGTTTGT 1518
...
NTU54774 1929 bp mRNA linear PLN 14-JAN-1997
Nicotiana tabacum glutamate decarboxylase (NtGAD1) mRNA, complete cds.
U54774.1 GI:1777920
Nicotiana tabacum (common tobacco)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
Dharmasiri,M.A.N., Lu,Y.T. and Harrington,H.M.
Cloning and sequencing of a tobacco cDNA encoding glutamate decarboxylase
Unpublished
2 (bases 1 to 1929)
Dharmasiri,M.A.N., Lu,Y.T. and Harrington,H.M.
Direct Submission
Submitted (11-APR-1996) M.A. Nihal Dharmasiri, Plant Molecular Physiology, University of Hawaii, 3190 Maile Way, Honolulu, HI 96822, USA
Location/Qualifiers
1..1929 /organism="Nicotiana tabacum"
/mol_type="mRNA"
/strain="Xanthi"
/db_xref="taxon:4097"
1..1929 /gene="NtGAD1"
66..1556 /note="calmodulin regulated enzyme; calmodulin-binding protein"
/codon_start=1
/product="glutamate decarboxylase"
/protein_id="AAB40608.1"
/db_xref="GI:1777921"
/translation="MWLSKTAASESDVSIHSTFASRYVTSLSRPFKMPENSIKKAAYQ IINDELMLDGNPRLNASFVITWMEPCNKLMMSDINKNYVDMDPEYPTVTELQNRVYN

```

```

MIAHLFNAPLGDGETAVGVGTGSSSEAIMLACLAFKRKQNMKMAQKFCDFKFNIVTG
ANYQVWEKFAFYFEVLEKAEVLSGYYVMDPEKAVMVDENITICVAAILKGSINVGSEF
EYKRLNDLILIEKNKTDPTPIHDAASGEFTAPFLYPELEWDFLPLPLKSNVNGSGH
KYLIVAGLIGWALWRNKEDLPDLFIHINVLGADQPTFLFNESKSGSSOVITAOYQILR
LGEFGYVWMEQENARVLRLEKSGRPNISKIEIIVPLVAFSLKDNSSQNEFFIIS
ETRRRGLIIPAYTFPFPNQRVTLRVLRVDFRFLAERLVLDIEKVLHELDITLFPAR
VNKLAVASRANGSGVHKHTDREVLQLEITTAWKRFVADKKKKTNGVC"
BASE COUNT 607 a 323 c 416 g 583 t
ORIGIN
Alignment Scores:
Pred. No.: 1.29e-179 Length: 1929
Score: 232.00 Matches: 427
Percent Similarity: 91.85% Conservative: 35
Best Local Similarity: 84.89% Mismatches: 33
Query Match: 85.35% Indels: 8
DB: Gaps: 3
US-10-006-852-2 (1-502) x NTU54774 (1-1929)
QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 66 ATGTTCTGTCCTCAAGACAGCGTCCGAAAGTGCCTCCATCCACTCCACTTCGTTCC 125
QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 126 CGATATGTTCTACTTCTCCGAGGTTTAAAGTCCAGAGAAATTCGATACCAAGGAA 185
QY 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGluValAsnProArgLeuAsnLeu 60
Db 186 GCAGCATCAAAATCAAAATGATGAGCTTATGTTAGTGGAAATCCAGACTAAATTA 245
QY 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
Db 246 GCATCTTTTGTGACACATGATGGAACAGAGTGTACAAACTGATGATGGATTCAT 395
QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
Db 306 AACAAAGAAATTAACCTTGACATGGATGAATACCCCTGTAAACCACTCAACTTCAG 365
QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
Db 366 GTAACATGATGACTCATTTGTTTAAACGACCACTTTGGAGATGGAGAGACTGCG 425
QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
Db 426 GTTGGAACTCTGGACTCTGAGGCTTATATCTTGTCTGGATTTCAAGAGAAA 485
QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
Db 486 TGGCAAAATATAAATGAAAGCCCAAGCCCTGTGACAAGCCCAATATFTGCACTGG 545
QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
Db 546 GCCAATGTCAGGTGTGGGAGAAATTTGCAAGGTAATTTTGAAGTGGAGCTAAAG 605
QY 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
Db 606 GTAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 665
QY 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
Db 666 GAGAACACAATTTGTAGTCTGCTATCTTGGGTTCCACTCTCAACTCAATGTTGAAT 725
QY 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
Db 726 GTTAAAGCGCTTGAATGACCTCTTGATTTGAGAGAAACAAGAAACCGGGTGGACAC 785
QY 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
Db 786 ATTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 845
QY 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280

```

JOURNAL Submitted (23-FEB-2001) Department of Plant Agriculture, University of Guelph, Division of Biotechnology, Guelph, Ontario N1G 2W1, Canada

FEATURES Location/Qualifiers

source 1..1672 /organism="Nicotiana tabacum" /mol_type="mRNA" /cultiivar="Samsun NN" /db_xref="taxon:4097" 4..1494 /EC_number="4.1.1.15" /function="converts glutamate to gamma-aminobutyric acid" /note="GAD; GAD3; NtGAD3; calcium/calmodulin-dependent enzyme" /codon_start=1 /product="glutamate decarboxylase isozyme 1" /protein_id="AAK18620.1" /gb_xref="GI:13310813" /translation="WVLSKTAASEDYSIHSFTASRYVRLSFRFKMPENSIPKEAAAY IINDELMDNPRMLASVFTWMEPECNKMMDSINKNYVDMGFFVFTTEIQNRCV MIAHLFNAPLGDGETAVGVTVGSEAIMLAGLAFKRWQKMKAGKPKCDKPNVTV ANVQWCFARYFEVELKEVLSGYYVMDPEKAVMVDENTICVAAILGSLTNGEF EDVKLNDLLEKKEBTGWDPIHVDAAAGGFIAFLYPELEWDRPLPLVKSINVSGH KYGLVYAGIWMALWRNKDLPDELIPHNYLGDQPTFTLNFSGKSSOVIAGYOLIR LGPEYKVMWENCQENARVLRGLKESGRFNLIISKEIGVPLVAFSLKONDSCHNEEIS EHIRFGWIIIPAYTTPFNQAQHVLRVIREDFSRILAEARLVIDLEKVLHLDLTPAR VNAKLAFAEANGSVHKKTDREVQLLEITAAWKKFVADKKKKTNGVC"

BASE COUNT 511 a 280 c 392 g 483 t 6 others

ORIGIN

Alignment Scores: 2.36e-179 Length: 1672

Prod. No.: 2228.00 Matches: 426

Score: 91.65% Conservative: 35

Percent Similarity: 84.69% Mismatches: 34

Best Local Similarity: 85.20% Indels: 8

Query Match: 8 Gaps: 3

DB:

US-10-006-852-2 (1-502) x AF352732 (1-1672)

QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20

Db 4 AUGGTTCTGTCCAGACAGCGTCCGAAAGTGCAGTCTCCATCCACTCCACTTCGGTTCC 63

QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40

Db 64 CGATATGTTCCAACTCTCTCCAGGTTAAAGATGCCAGAAATTCATACCAAAAGAA 123

QY 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60

Db 124 GCAGCATATAAATCAATAATGATGAGCTTATGTTAGATGGAAATCCAAAGGCTAAATTA 183

QY 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80

Db 184 GCATCTTTTGTGACACATGATGGAAACCCAGAGTGTAAATAGTTAATGATGGATCCAT 243

QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100

Db 244 AACCAAGAACTACCTTGCATGGTGGTGAATACCCCTGTAACCCAGCTGAGCTTCAGAATCGATG 303

QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120

Db 304 GTAATAATGATGACTCATTTGTTTAAACCCACCTTGGAGATGGAGAGACTGCAGTTGGA 363

QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140

Db 364 GTTGGAACTGTGGATCCTTGAAGGCTTATGCTTGTGGATTAGCTTCAAGAAAAA 423

QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160

Db 424 TGGCAAAATAAATAAGAAAGCCCAAGGAAAGCCCTGGCAAGCCCAATAATTTGCTACTGGT 483

QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180

JOURNAL Submitted (23-FEB-2001) Department of Plant Agriculture, University of Guelph, Division of Biotechnology, Guelph, Ontario N1G 2W1, Canada

FEATURES Location/Qualifiers

source 1..1672 /organism="Nicotiana tabacum" /mol_type="mRNA" /cultiivar="Samsun NN" /db_xref="taxon:4097" 4..1494 /EC_number="4.1.1.15" /function="converts glutamate to gamma-aminobutyric acid" /note="GAD; GAD3; NtGAD3; calcium/calmodulin-dependent enzyme" /codon_start=1 /product="glutamate decarboxylase isozyme 1" /protein_id="AAK18620.1" /gb_xref="GI:13310813" /translation="WVLSKTAASEDYSIHSFTASRYVRLSFRFKMPENSIPKEAAAY IINDELMDNPRMLASVFTWMEPECNKMMDSINKNYVDMGFFVFTTEIQNRCV MIAHLFNAPLGDGETAVGVTVGSEAIMLAGLAFKRWQKMKAGKPKCDKPNVTV ANVQWCFARYFEVELKEVLSGYYVMDPEKAVMVDENTICVAAILGSLTNGEF EDVKLNDLLEKKEBTGWDPIHVDAAAGGFIAFLYPELEWDRPLPLVKSINVSGH KYGLVYAGIWMALWRNKDLPDELIPHNYLGDQPTFTLNFSGKSSOVIAGYOLIR LGPEYKVMWENCQENARVLRGLKESGRFNLIISKEIGVPLVAFSLKONDSCHNEEIS EHIRFGWIIIPAYTTPFNQAQHVLRVIREDFSRILAEARLVIDLEKVLHLDLTPAR VNAKLAFAEANGSVHKKTDREVQLLEITAAWKKFVADKKKKTNGVC"

BASE COUNT 511 a 280 c 392 g 483 t 6 others

ORIGIN

Alignment Scores: 2.36e-179 Length: 1672

Prod. No.: 2228.00 Matches: 426

Score: 91.65% Conservative: 35

Percent Similarity: 84.69% Mismatches: 34

Best Local Similarity: 85.20% Indels: 8

Query Match: 8 Gaps: 3

DB:

US-10-006-852-2 (1-502) x AF352732 (1-1672)

QY 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20

Db 4 AUGGTTCTGTCCAGACAGCGTCCGAAAGTGCAGTCTCCATCCACTCCACTTCGGTTCC 63

QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40

Db 64 CGATATGTTCCAACTCTCTCCAGGTTAAAGATGCCAGAAATTCATACCAAAAGAA 123

QY 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60

Db 124 GCAGCATATAAATCAATAATGATGAGCTTATGTTAGATGGAAATCCAAAGGCTAAATTA 183

QY 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80

Db 184 GCATCTTTTGTGACACATGATGGAAACCCAGAGTGTAAATAGTTAATGATGGATCCAT 243

QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100

Db 244 AACCAAGAACTACCTTGCATGGTGGTGAATACCCCTGTAACCCAGCTGAGCTTCAGAATCGATG 303

QY 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120

Db 304 GTAATAATGATGACTCATTTGTTTAAACCCACCTTGGAGATGGAGAGACTGCAGTTGGA 363

QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140

Db 364 GTTGGAACTGTGGATCCTTGAAGGCTTATGCTTGTGGATTAGCTTCAAGAAAAA 423

QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160

Db 424 TGGCAAAATAAATAAGAAAGCCCAAGGAAAGCCCTGGCAAGCCCAATAATTTGCTACTGGT 483

QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180

846 TGGGACATTTAGATGGCCATTTGGTGAAGAGATTAATTAACCTGAGTGGTCAACAATATGGTCTT 905

281 ValTyrAlaGlyIleGlyTyrValIleTyrArgAsnLysGluAspLeuProGluGluLeu 300

906 GTTATGCTGGTATTTGGTGGCCATTTGGAGGAAVAGGAAGACTTACCTGACGAACTT 965

301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320

966 ATCTTCCACATTAATTAATTTGGTGGTCACTCAACCTACTTCACTCAACTCTCTAAA 1025

321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340

1026 GGTTCATAGCCCAAGTAATGCTCAATATTAACAACCTTATTCCTGGGTTTGGAGGTTAC 1085

341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360

1086 AAGAAATGTTATGGAGATTTGCAAGAAAATCCRAAGGCTACTAAGAAAGGACTTGA 1145

361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380

1146 AGTGAAGATTAACAATAATATCAAAAGAAATTTGGAGTTCCATTTAGTAGCTTTCTCT 1205

381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgAspTyrGlyTyr 400

1206 AAGACAACTGCAACATGAGTTGAAATTTCTGAAATTTCTGAACTTTAAGAAATTTGG 1265

401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420

1266 ATTATCTCTGATATATGTCACCAAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1325

421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440

1326 ATTAGAAAGATTTCTCCGTPACACTCCCGAGGACTGTTAATAGACATTTGA 1385

441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460

1386 CTCACAGGCTAGACACACTTCCGGGAGGTCACCGTAGCTAGTCGGCGT----- 1436

461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480

1437 ---GCCGAGGGAATGGCAGCGGCTGCAT-----AAGAAACAGATAGAAAGTG 1484

481 GlnArgAspIleIleThrGlyTyrLysLysPheValAlaAsp---ArgLysLysThrSer 499

1485 CAGCTTGAGATTACTACTGTCATGGAAGAAATTTGTTGCTGATAAAGAAAGAACTAAC 1544

500 GlyTyleCys 502

1545 GGAGTTTGT 1553

SULT 11

352732

CUS 1672 bp mRNA linear PLN 02-MAY-2003

FINITION Nicotiana tabacum glutamate decarboxylase isozyme 1 mRNA, complete cds.

CESSION AF352732.1 GI:13310812

RSTON

YWORDS Nicotiana tabacum (common tobacco)

ORGANISM Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 1672) Yevtushenko, D., McLean, M.D., Peiris, S.E., Van Cauwenbergh, O.R. and Shelp, B.J.

AUTHORS Two isoforms of tobacco glutamate decarboxylase are regulated by calcium/calmodulin and differ in organ distribution

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1672) Yevtushenko, D., McLean, M.D. and Shelp, B.J.

AUTHORS Direct Submission

TITLE

```

484 GCCAATGTCAGGTTGTGGGAAAATTTCCAAAGGTATTTTGAAGTGGAGTTGAAAGAA 543
181 VallyLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
544 GTAAAATTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 603
201 GlnAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
604 GAGAATACAAATTTGTAGCTGCTATCTTGGGTTCCACTCTCAATGTTGTAATTTGAAGAT 663
221 VallyLeuLeuAsnAspLeuValGlnLysAsnLysGluThrGlyTyrAspThrPro 240
664 GTTAAAGCCCTTGAATGACCTCTGATGATGAGAAGAACAAAGAAAACCGGTGGGACATCCA 723
241 IleHisValAspAlaIleSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
724 ATTCAATGCGATGCAGCAAGTGTGGATTTCATTGCACCAATTCCTTATCCGTGAGCTTGA 783
261 TrpAspPheArgLeuProLeuValIleSerIleAsnValSerGlyHisLysTyrGlyLeu 280
784 TGGGATTTTAGATTAACCAATGCTGAAGAGTATTAATGTGAGTGTGCACAAATATGGTCTT 843
281 ValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluLeu 300
844 GTCTATGCTGTTATTTGGTGGCCATTTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 903
301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
904 ATTTTCCACATTAATTAATCTTGGTGTGATCAACTACTTTCACCTCAACTTCTCTAAA 963
321 GlySerSerGlnValIleAlaGlnTyrTyrGlnIleLeuIleArgLeuGlyHisGlyTyr 340
964 GGTTCAGCCCAAGTAATTTGCTCAATPATTACCAACTATTCCTGCTGGGTTTGGAGGTTAC 1023
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValIleuArgGluGlyLeuGluLys 360
1024 AAGAAATGTTATGGAGAAATTCGAAGAAATCAAGGGTACTAAGAAAGGACTTGAATAA 1083
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
1084 AGTGAAGATTCATATTAATCTCCAAAGAAATTTGGAGTTCCAGTTCCATAGTAGCTTCTCT 1143
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTyr 400
1144 AAAGACACAGTCAACAAATGATGTTCCAAAATTTCTGAAACTTTAGAAAGATTTGGATGG 1203
401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValIleuArgValVal 420
1204 ATTAATCTGCATATACATGCTCCAAATGCTCAACATGCTCAACATGCTCAGAGTTGTT 1263
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
1264 ATTAGAAGATTTCTCCCTACACTCCGCGAGGACTGGTGTATAGACATTTGAAAAGTC 1323
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1324 CTCCACGAGCTAGACACACTTCCGCGAGGCTCAACGCTAAGCTGCGCGTG----- 1374
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1375 ---GCCGAGCCGAAATGGCAGCGGCGTGTGAT-----AGAAACAGATAGAGAGTG 1422
481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
1423 CAGCTGGAGATTAAGTCTGATGGAAGAAAATTTGTTGCTGATAGAAAGAAAGACTAAT 1482
500 GlyIleCys 502
1483 GGAGTTTGT 1491

```

```

LOCUS AX654603 1479 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 4473 from Patent WO03000898.
ACCESSION AX654603
VERSION AX654603.1 GI:29157417
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Kato, Y., F., Quan, S., Tao, Y., Whitcham, S., Xie, Z., Zhu, T. and Zou, G.
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 4473 03-JAN-2003;
Synygenta Participations AG (CH)
FEATURES
Location/Qualifiers
1..1479
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"
BASE COUNT 298 a 478 c 472 g 231 t
ORIGIN
Alignment Scores: 4.77e-177 Length: 1479
Pred. No.: 2200.00 Matches: 420
Score: 2200.00 Conservative: 35
Percent Similarity: 90.64% Mismatches: 37
Best Local Similarity: 83.67% Indels: 10
Query Match: 84.13% Gaps: 2
DB:
US-10-006-852-2 (1-502) x AX654603 (1-1479)
QY 1 MetValIleuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 1 ATGGTCTCTCCAAAGCCGCTCCCGAGGTGACATGTCGTGCACCTCCACCTTCGGCTCC 60
QY 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 61 CGCTACGTCGCGCCCTCCCTCCCAAGTACCAGTACCAGTACCAGTACCAGTACCAGTACC 120
QY 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 121 GCGCGGTACAGATCATCAACGACGAGCTGATGCTGGACGGCAACCCCGCTGAACTG 180
QY 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
Db 181 GCGTCTGTCACCACTGGATGGAGCCGAGTGGACAAAGCTCATATGGCCGCGCACT 240
QY 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
Db 241 AACAAAGAACTACGTCGACATGGACGAGTACCCCGTCCACCAGCTCCAGAACCCGGTGC 300
QY 101 ValAsnMetIleAlaHisIleuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
Db 301 GTGAACATGATCGCCACCTGTTCCACCGCCCTCGGGAGGACGAGACGGCGGTGGGC 360
QY 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
Db 361 GTGGCACGCTGGGTTCGTCGAGGCACTATGCTGGCCGGGCTGGCCCTCAAGCCGCGG 420
QY 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
Db 421 TGGCAGAAACAAGCCCAAGGCGGAGGAAAGCCGTTCCGAAAGCCCAACATCATCCCGGC 480
QY 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGlnValGluLeuLysGlu 180
Db 481 GCCAACGTGCAGGTGTGGGAGAGTTCGCCCGCTACTTCGAGGTGGAGCTCAAGGAG 540
QY 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnAlaValAspMetValAsp 200
Db 541 GTGAAGCTCCGACGGCTACTACTGCTGAGTGGAGCCCGGAGAGCCCGCTGACATGTTCAAC 600

```


201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
601 GAGAACCACTCGCTCGCGCCATCTCGCTCCACCCCTCAACCGCGAGTTCGAGGAC 660
221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
661 GTC AAGCTACTCAACGACCTCTCGCAAGAAAGAACGAGGACTGGTGGAGACGCCG 720
241 IleHisValAspAlaIleSerGlyGlyPheIleAlaProPheLeuTrpProGluLeuGlu 260
721 ATCCACGTTGAGCGGGGAGCGGGTTCATCGCGCGTTCCTGTATCCCGGAGCTGG 780
261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTrpGlyLeu 280
781 TGGGACTTCGGCTCGCGTGGTGAAGAGCATCAACCTGAGCGGTCAAGTACCGGCTC 840
281 ValTrpAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluLeu 300
841 GTCVACCGCGGATCGGCTGGTGTGTCATCTGGGCGCAACGAGGAGGACCTGCGG 900
301 IlePheHisIleAsnTrpLeuGlyValAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
901 ATCTTCCACATCACTACTCGGCACCGACCGACCAACCTTCCCTCACTTCTCCAG 960
321 GlySerSerGlnValIleAlaGlyTrpGlnLeuIleArgLeuGlyHisGluGlyTrp 340
961 GGCTCCAGCCAGGTCATCGCCCTACTACAGCTATCCAGCTATCCGCCACCGGCTC 1020
341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
1021 AGGAACATCATGGAGACTGCCACGAGAACCGGATGGTGTGAAGGAGGCTGGTGA 1080
361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
1081 ACGGGAGTTCGACATGCTGCCAAGGACGAGGGTGGCGGCTGGTGGCTTCTCGCTC 1140
381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTrpGlyTrp 400
1141 AAGGACCGGAGCGGACGAGGAGTTCGAGATCTCCGACATGCTGCGCGCTTCCGGTGG 1200
401 IleValProAlaTrpMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
1201 ATCGTGGCGGATACCATGCGCCCGCCAGCCGACGACGTCACGGTCTCCGGTGGCT 1260
421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
1261 ATCCGGAGGAGTTCAGCGCACCTCCCGGAGCGCTCGCTCCGACATCCAGAGGTTG 1320
441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
1321 ATGTACCAGCTCGACGCGCTCCCTCCAGGCTCATGCCCCCGGCGCGCGCGCGCG 1377
461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
1378 -----CTGCTGGTGGTCCCAAGAAAGTTCGAGCTCGAGACG 1413
481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
1414 CAGCGTGGTGGAGCGGCTGGGAAAGTTCGTGCTC---GCCAAGAGGACCAACGGC 1470
501 IleCys 502
1471 GTCTCC 1476

source
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Halminori"
/db_xref="taxon:39947"
1. .1502
21. .1499
/genes="GAD3"
/BC_number="4.1.1.15"
/codon_start=1
/product="glutamate decarboxylase"
/protein_id="AA059316.1"
/db_xref="GI:28911953"
/translation="MVLKAVSESDMSVHSTFASVVRASLPRVRMPENSIKPEAAVO
ANVQVWEKFPARYFEVLEKVKLRDGYVMDEKAVDMVNTI CVAAILGSLTNGEF
EDVLIINDLIDKKNKETMETPIHVDAASGGFIAPFLYPELWDFRLPWKGSINYSGH
KGYVYAGIWCITWRNKEDPELIFHNYLGTDDPTFLNFSKGSOSVIAQVYOLIR
KGFEGVYIWNCHENAMVLEKLVKGTGFDIVSKDEGVLVAFSLKDRSRHDEFEIS
DMRRFRWLVAYTMDPPDAQHVTVLKVIREFEFRLAERLVLIDIEKVMYQLDALLPSR
LMPPVPPALVLYVAKKSELETQRSVTEAWKFKVLAKRTNGVC"
BASE COUNT 303 a 483 c 479 g 237 t
ORIGIN
Alignment Scores: 4.86e-177 Length: 1502
Pred. No.: 2200.00 Matches: 420
Score: 2200.00
Percent Similarity: 90.64% Conservative: 35
Best Local Similarity: 83.67% Mismatches: 37
Query Match: 84.13% Indels: 10
DB: Gaps: 2
US-10-006-852-2 (1-502) x AY187941 (1-1502)
Qy 1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
Db 21 ATGCTCTCTTCAAGCGCGTCTCCGAGAGTGACATGTCGCTGACCTCCACCTTCGCTCC 80
Qy 21 ArgTrpValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
Db 81 CGCTACGTCGCGCGCTCCCTCCCAAGGATCCGAGTACCGGAGAACTCGATCCCGAAGGAG 140
Qy 41 AlaAlaTrpGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
Db 141 GCGGCGTACCAAGATCAACGACGAGCTGATGCTGGACGGCAACCCCGGCTGAACCTG 200
Qy 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerIle 80
Db 201 GCTCGTTCGTCCACCGCTGGAGCCCGAGTGGCAAGCTCATATGGCGCCATC 260
Qy 81 AsnLysAsnTrpValAspMetAspGluTrpProValThrGluLeuGlnAsnArgCys 100
Db 261 AACAGAACTACGTCGACATGGAGTACCCCGTCCACCAGGCTCCAGAACCCGGTGC 320
Qy 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 1502)
Liu,L.L., Zhai,H.Q. and Wan,J.M.
New Rice GAD gene GAD3
Unpublished
2 (bases 1 to 1502)
Liu,L.L., Zhai,H.Q. and Wan,J.M.
Direct Submission
Submitted (27-NOV-2002) State Key Laboratory of Crop Genetics
Germplasm Enhancement, Jiangsu Plant Gene Engineering Research
Center, Weizang, Nanjing, Jiangsu 210095, P. R. China
Location/Qualifiers
1. .1502
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Halminori"
/db_xref="taxon:39947"
1. .1502
21. .1499
/genes="GAD3"
/BC_number="4.1.1.15"
/codon_start=1
/product="glutamate decarboxylase"
/protein_id="AA059316.1"
/db_xref="GI:28911953"
/translation="MVLKAVSESDMSVHSTFASVVRASLPRVRMPENSIKPEAAVO
ANVQVWEKFPARYFEVLEKVKLRDGYVMDEKAVDMVNTI CVAAILGSLTNGEF
EDVLIINDLIDKKNKETMETPIHVDAASGGFIAPFLYPELWDFRLPWKGSINYSGH
KGYVYAGIWCITWRNKEDPELIFHNYLGTDDPTFLNFSKGSOSVIAQVYOLIR
KGFEGVYIWNCHENAMVLEKLVKGTGFDIVSKDEGVLVAFSLKDRSRHDEFEIS
DMRRFRWLVAYTMDPPDAQHVTVLKVIREFEFRLAERLVLIDIEKVMYQLDALLPSR
LMPPVPPALVLYVAKKSELETQRSVTEAWKFKVLAKRTNGVC"

gene
CDS

RESULT 13
OCUS AY187941
DEFINITION Oryza sativa (japonica cultivar-group) glutamate decarboxylase
(GAD3) mRNA, complete cds.
-CESSION AY187941
VERSION AY187941.1 GI:28911952
KEYWORDS

67 ATGGTCTGTCACAGACAGCGTCCGAAAGTGGAGTTCCTCCCTAGTAGCAATTTTCCTT 1206
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 127 CGATVAGTTCGAATCTCTCCAGGTTTAAATGCCAGAGAAATTCATACCAAGGA 186
 41 AlaAlaTyrGlnIleLeuAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 187 GCAGCATATCAGATATAAATGATGAGCTTATGTAGATGAAATCCAAAGCTAAATTA 246
 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 247 GCATCTTTCGTTACACATGATGGAGCCAGAAATGTAATACGTTAATGATGGATCCAIT 306
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
 307 AACAAAGTACGTTGACATGGATGATACCCCTGTAACCACTGAGCTTCAGAAATCGATGT 366
 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluAlaGluThrAlaValGly 120
 367 GTAATAATGATGATGCTTCTTAAATGCAACCACTTGGAGATGGAGAGACTGCGATTGGA 426
 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 427 GTTGGAACTGTTGGATCCTCTGAAAGCTATTATGCTTCTGATGAGCTTAAAGAGAAA 486
 141 TrpGlnAsnLysArgLysAlaGluGlyProValAspLysProAsnIleValThrGly 160
 487 TGGCAAAATAAAGAAAGCCCAAGCAAGCCCTTTGATAAGCCCAATAATGTCACCGGT 546
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 547 GCTAATGTCAGAGTGTGGGAGAAATTTGCAAGGTATTTGAAAGTGGAGTTGAAAGAA 606
 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnAlaValAspMetValAsp 200
 607 GTAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 667 GAGAAATACCAATTTGCTGCTATCTTAGTTCAACACTCAATGTTGAATTTGAAGAT 726
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 727 GTTAAAGCGTTGAAATGACCTTTGATGAGAAAGCAAGAAACCGGGTGGACACTCCA 786
 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 787 ATTCATGTGGATGACGACAGTGGTGGATTTATGACCACTTCTTTATCCAGAGCTTGA 846
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 847 TGGACTTTAGATTGCCATTCGAGAGAGTATTAAATGATGATGATGATGATGATGATGATGAT 906
 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 907 GTCATGCTGTTGTTGGCCATTTGGAGAAATAGGAAGACTTTCCTGATGAACTT 966
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 967 ATTTCCACATCAATACCTGCTGATCAACCTACTTTCCTCACTCACTTCTCTAAA 1026
 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnIleLeuIleArgLeuGlyHisGluGlyTyr 340
 1027 GGTTCAGCAAGTAAATGCTCAATATACCACCTTATTCGCTTGGTTTTGAGGGTTAC 1086
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
 1087 AAGAATGTTATGGAGAAATGTCAGAAATGCCAAGGATTAAGAGAGAAATGAAAAA 1146
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380

1147 AGTGAAGATTCAACATAAATCTCCAAAGAAATTTGGAGTTCCTCCCTAGTAGCAATTTTCCTT 1206
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
 1207 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTTTAGAAATTTGGATGG 1266
 401 IleValProAlaTyrThrMetProAsnAlaGlnHisIleThrValLeuArgValVal 420
 1267 ATTTGTTTGGCATATACCTATGTCACCAAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1326
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 1327 ATTAGAAGATTTCTCCGGCACACTAGCGGAGCCTGGTAAATAGACATTTGAAAAGATC 1386
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 1387 TTCCACGGAGTAGACACACTTCGGCGAGGCTCAACGGCTAAGCTAGCCGTG----- 1437
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 1438 ---GCCGAGGCAATGGCAGCGCGTGCAT-----AAGAAAACAGATAGAGAGTGT 1485
 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
 1486 CAGCTAGAGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1545
 500 GlyIleCys 502
 1546 GGAGTTTGT 1554

RESULT 15
 AF353615
 LOCUS AF353615.1 GI:13752461
 DEFINITION Nicotiana tabacum glutamate decarboxylase isozyme 3 mRNA, complete cds.
 ACCESSION AF353615
 VERSION AF353615.1
 KEYWORDS Nicotiana tabacum (common tobacco)
 SOURCE Nicotiana tabacum
 ORGANISM Nicotiana tabacum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 1776)
 AUTHORS Yevtushenko, D., McLean, M.D., Peiris, S.E., Van Cauwenbergh, O.R. and Shelp, B.J.
 TITLE Two isoforms of tobacco glutamate decarboxylase are regulated by calcium/calmodulin and differ in organ distribution
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1776)
 AUTHORS McLean, M.D., Yevtushenko, D. and Shelp, B.J.
 TITLE Direct Submission
 REFERENCE 3 (bases 1 to 1776)
 AUTHORS Submitted (26-FEB-2001) Department of Plant Agriculture, University of Guelph, Division of Biotechnology, Bovey Building, Guelph, Ontario N1G 2W1, Canada
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..1776
 /organism="Nicotiana tabacum"
 /mol_type="mRNA"
 /db_xref="taxon:4097"
 CDS 89..1564
 /EC_number="4.1.1.15"
 /function="converts glutamate to gamma-aminobutyric acid"
 /note="GAD; GAD4; NtGAD4"
 /codon_start=1
 /product="glutamate decarboxylase isozyme 3"
 /protein_id="AAK38667.1"
 /db_xref="GI:13752462"
 /translation="MVLSTKSESDVSVHSTFASVYVTSLSRFRWAENSIPKRAAFQ IINDELMDGNRFLNLSAFVITWMEPCDKLMMDSINKNYVDMDEYFTTELQRRCV N IARLNFAPLEEKTAIVLGLAFKRNQKRAEGRFYNKPNLIVG ANVQVCKEFANYFVELKREGYKMPDQVQVEMVDENTICVAAILGSLNGEFD EDVKLLDILLIEKNTGWNTPIHVDAASGGFIAPFLYPELEWDFRFLVKSINVSGH

KYGLVIVAGIGWVIRWTKQDIPBELEIPIHINYLQADQFTFFLNFSKSGSSQVIAQYQLIR
 LGYEVNVMENCRNAI VLRBGLKREINIVS KDEGVPVLAFAFKDNRSHRNEFEVS
 ETLLRFWIVPVTMPADAOHVLRVIRVDFSRLEAERLVLDIKVLVHLDLTLPAR
 LSAKLEVKLVKNGKFKPELEVRQEVNINWYKFKVLAARKAPVC"
 SE COUNT 520 a 343 c 418 g 491 t 4 others
 IGIN

Alignment Scores:
 ed. No.: 8.32e-176 Length: 1776
 ore: 2186.50 Matches: 421
 recent Similarity: 90.32% Conservative: 27
 st Local Similarity: 84.88% Mismatches: 39
 ery Match: 83.61% Indels: 9
 Gaps: 8 1

-10-006-852-2 (1-502) x AF353615 (1-1776)

1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
 |||
 89 ATGGTTCTCTCGAANAACCTCTCGATCGGACGTTTCGGTACACTCCACTTTGGCTCT 148
 |||
 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
 |||
 149 CGCTATGTTGGAATCTCCCTTCCAAGSTTTGAGATGCGGGRGAAATTCGATACCAAAAGAG 208
 |||
 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
 |||
 209 CCGGCAATTTCAAAATAAATAACGACGAGTTGATGTTGACGGGAATCCAAAGGCTGAACTTG 268
 |||
 61 AlaSerPheValThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
 |||
 269 GCTTCATTTGACAAATGATGGACGACGAGTGTGATAGCTTATGATGGACTCCATT 328
 |||
 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrGluLeuGlnAsnArgCys 100
 |||
 329 AACAAAGAACTATGTTGACATGATGATGAAATACCCTCTGTTACCACTGAGCTTCAGAAATCGCTGC 388
 |||
 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
 |||
 389 GTGAACATGATAGCAGCTTTTATTCAACCGCGCCACTAGAAAGAAAGGAGACAGCAGTTGGA 448
 |||
 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
 |||
 449 GTGGTACAGTTGGTTATCGGAGCCATAATGCTAGCGGGCTTGCAATCAAGAAAT 508
 |||
 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
 |||
 509 TGGCAAAAACAAACGCAAGCTGAGGCCAAACCTTACAATAAGCCCAACATTTGTCACCTGGC 568
 |||
 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
 |||
 569 GCCAATGTTCAAGTGTGCTGGGAAATTTGCCAACTATTTGAAAGTGGAAATGAAAGAA 628
 |||
 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnAlaValAspMetValAsp 200
 |||
 629 GTAAGCTTAAGGAAAGGTACTATGATGATGACCCAGCTCCAGGCTGTGGATGGTTGAT 688
 |||
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
 |||
 689 GAGAACACCAATTTGGTGTCTGCAACTCTGGGTTCAACCCCTTAATGGAGAAATTTGAAAGAT 748
 |||
 221 ValLysLeuLeuAsnAspLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
 |||
 749 GTCAAAGCTCTTGAATGATCTTTTGTGATGAAAGAAACAAAGCAACTGGATGAAACACCA 808
 |||
 241 IleHisValAspAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
 |||
 809 AITCATGTGGATGCAGCAAGTGGTGGATTTCAATGACCCTTCTCTGATCCCGAGACTGGAG 868
 |||
 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
 |||
 869 TGGGACTTTAGGCTTCCCTTAGTGAAGACATAAATGATGAGTGGGCACAAATATGGGCTT 928

281 ValTyrAlaGlyIleGlyTyrValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
 |||
 929 GTCTATGCTGTTGTTGGTTATTTGGAGGACCAACAAGATTTGCCTCAGAAATC 988
 |||
 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
 |||
 989 AITTTCCACATCAACTACTCTGGAGCTGATCAGCCACTTTACTCTCAATTTCTCCAAA 1048
 |||
 321 GlySerSerGlnValIleAlaGlnTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
 |||
 1049 GGTCAAGTCAAGTCAATGCTCAATATTATCAGCTAAATCCGTTTGGGCTATGAGGGGTAC 1108
 |||
 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuLeuArgGluGlyLeuGluLys 360
 |||
 1109 CGAAATGTAATGGAGAACTGTCGGAAAATGCCAATTTGCTAAGAAAGAACTCGAAAAA 1168
 |||
 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
 |||
 1169 ACAGACAGTTTTTACATAGTCTCCAAAGATGAGAGGTTCCTCTTGGTGGCCTTTCCCTC 1228
 |||
 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgTyrGlyTrp 400
 |||
 1229 AAGGACAAATAGCCGTCACAAATGAGTTGAGGRTCCGAGACGCTCCGTAGGTTCCGGTGG 1288
 |||
 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
 |||
 1289 ATCGTCCCGCCCTACACGATGCCGCTGACGCCCAACACGTCACCGGTGCTTCGTGTGGTG 1348
 |||
 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
 |||
 1349 ATCCGGGAGACTTCTCGGAAACCTTAGCAGACGCTCTCGTCTCGACATTTGTCAGGGTC 1408
 |||
 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
 |||
 1409 CTCACGAGCTGGACACACTTCCAGTAGGCTGAGCGCCAAATTA----- 1453
 |||
 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
 |||
 1454 -----GAGGAGTGAAGCTGGTCAAGAAATGGAAAGAAATTTGAACTTGAAGTT 1501
 |||
 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLys 496
 |||
 1502 CAAAGGAAATTTACCAATTTATTTGGAAGAAATTTGTTTATAGCTAGGAAA 1549
 |||

Search completed: October 22, 2003, 16:47:12
 Job time : 4597 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

1 protein - protein search, using sw model

Search time 443 Seconds
(Without alignments)
189.764 Million cell updates/sec

US-10-006-852-2

Sequence: I MVLSHAVSESDVSHSTFAS.....DIITGWKFKVADRKKTSGIC 502

oring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

629382 seqs, 167460630 residues

tal number of hits satisfying chosen parameters: 629382

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: result No., Score, Match, Length, DB, ID, Description. Contains 18 rows of application summaries.

Table with columns: 16-45, 393.5-101.5, 15.0-3.9, 191-825, 11-9, US-09-989-442-145, Sequence 145, App, Sequence 8, Appl, Sequence 2, Appl, Sequence 8, Appl, Sequence 16, Appl, Sequence 11, Appl, Sequence 6, Appl, Sequence 4, Appl, Sequence 6, Appl, Sequence 6, Appl, Sequence 2, Appl, Sequence 4, Appl, Sequence 4, Appl, Sequence 4, Appl, Sequence 4, Appl, Sequence 18, Appl, Sequence 10, Appl, Sequence 10, Appl, Sequence 10, Appl, Sequence 20, Appl, Sequence 7900, Ap, Sequence 74, Appl, Sequence 77, Appl, Sequence 167, App, Sequence 2, Appl, Sequence 11741, A, Sequence 5882, Ap, Sequence 13079, A

ALIGNMENTS

RESULT 1
; Sequence 2, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
; APPLICANT: Kinnerseley, Alan M.
; APPLICANT: Turano, Frank J.
; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
; FILE REFERENCE: 7224-65
; CURRENT APPLICATION NUMBER: US/10/006,852
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/246,367
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-006-852-2

Query Match 100.0%; Score 2615; DB 15; Length 502;
Best local Similarity 100.0%; Pred. No. 8.9e-245;
Matches 502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVLSHAVSESDVSHSTFASRYVRTSLPRPKPENSIPKEAAVQIINDEMLDGNPRLNL 60
Db 1 MVLSHAVSESDVSHSTFASRYVRTSLPRPKPENSIPKEAAVQIINDEMLDGNPRLNL 60
Qy 61 ASFVTTWMEPECDKLI MSSI NKYVDMDEYPTVTELQNR CNVMIAHLFNAPLEEAETAVG 120
Db 61 ASFVTTWMEPECDKLI MSSI NKYVDMDEYPTVTELQNR CNVMIAHLFNAPLEEAETAVG 120
Qy 121 VGTVGSSEALMAGLAFKRWQNKRAEKQPKVDPKPNIVTGANVQVCWEKPFARYFVELKE 180
Db 121 VGTVGSSEALMAGLAFKRWQNKRAEKQPKVDPKPNIVTGANVQVCWEKPFARYFVELKE 180
Qy 181 VKLSEGYVMDPQAVDMVDENTICVADILGSLNGSEFEDVKLLNDLLVEKNKETGMDIF 240

RESULT 4
3-10-005-602-2
Sequence 2, Application US/10005602
Publication No. US20030110530A1
GENERAL INFORMATION:
APPLICANT: Barry Shelp
APPLICANT: Alan Bown
TITLE OF INVENTION: TRANSGENIC PLANTS HAVING REDUCED
FILE OF INVENTION: SUSCEPTIBILITY TO INVERTEBRATE PESTS
FILE REFERENCE: P8AUS3
CURRENT APPLICATION NUMBER: US/10/005,602
CURRENT FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 496
TYPE: PRT
ORGANISM: tobacco plant
-10-005-602-2

Query Match 85.1%; Score 2226; DB 15; Length 496;
Best Local Similarity 84.7%; Pred. No. 4.9e-207;
Matches 426; Conservative 35; Mismatches 34; Indels 8; Gaps 3;

1 MVLHVSSEDSVSHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
1 MVLKSTASESDVSIHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
61 ASFVTTWMEPECDKLIIMSSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
61 ASFVTTWMEPECNTLMWDSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
181 VKLSEGYVMDPQQAQVDMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
181 VKLSDGYVMDPEKAVEMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
361 TERFNIVSKDEGVPLVAFSLKSDSSCHTEFEISDMLRRYGVIPAYTPPNAOHITVLRVY 420
361 SGRFNIIKEIGVPLVAFSLKSDNSOHNEFEISLRRFRGWIIVLAYTPPNAOHITVLRVY 420
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480

481 QRDITGWKKFVAD-RKKTSGIC 502
474 QLEITTAWLKRVADKXKKTNGVC 496

RESULT 5
3-10-006-852-14
Sequence 14, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
APPLICANT: Kinnerseley, Alan M.
APPLICANT: Turano, Frank J.
TITLE OF INVENTION: Methods for Regulating Plant GABA Production
FILE REFERENCE: 7224-65
CURRENT APPLICATION NUMBER: US/10/006,852
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: US 60/246,367
PRIOR FILING DATE: 2000-11-07
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent in version 3.1
SEQ ID NO 8
LENGTH: 493
TYPE: PRT
ORGANISM: Arabidopsis thaliana
CURRENT APPLICATION NUMBER: US/10/006,852

Query Match 84.1%; Score 2199; DB 15; Length 496;
Best Local Similarity 84.1%; Pred. No. 2e-204;
Matches 423; Conservative 34; Mismatches 38; Indels 8; Gaps 3;

1 MVLHVSSEDSVSHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
1 MVLKSTASESDVSHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
61 ASFVTTWMEPECDKLIIMSSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
61 ASFVTTWMEPECNTLMWDSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
181 VKLSEGYVMDPQQAQVDMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
181 VKLSDGYVMDPEKAVEMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
361 TERFNIVSKDEGVPLVAFSLKSDSSCHTEFEISDMLRRYGVIPAYTPPNAOHITVLRVY 420
361 SGRFNIIKEIGVPLVAFSLKSDNSOHNEFEISLRRFRGWIIVLAYTPPNAOHITVLRVY 420
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480

481 QRDITGWKKFVAD-RKKTSGIC 502
474 QLEITTAWLKRVADKXKKTNGVC 496

RESULT 6
US-10-006-852-8
Sequence 8, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
APPLICANT: Kinnerseley, Alan M.
APPLICANT: Turano, Frank J.
TITLE OF INVENTION: Methods for Regulating Plant GABA Production
FILE REFERENCE: 7224-65
CURRENT APPLICATION NUMBER: US/10/006,852
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: US 60/246,367
PRIOR FILING DATE: 2000-11-07
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent in version 3.1
SEQ ID NO 8
LENGTH: 493
TYPE: PRT
ORGANISM: Arabidopsis thaliana
CURRENT APPLICATION NUMBER: US/10/006,852-8

Query Match 84.1%; Score 2199; DB 15; Length 496;
Best Local Similarity 84.1%; Pred. No. 2e-204;
Matches 423; Conservative 34; Mismatches 38; Indels 8; Gaps 3;

1 MVLHVSSEDSVSHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
1 MVLKSTASESDVSHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
61 ASFVTTWMEPECDKLIIMSSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
61 ASFVTTWMEPECNTLMWDSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
181 VKLSEGYVMDPQQAQVDMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
181 VKLSDGYVMDPEKAVEMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
361 TERFNIVSKDEGVPLVAFSLKSDSSCHTEFEISDMLRRYGVIPAYTPPNAOHITVLRVY 420
361 SGRFNIIKEIGVPLVAFSLKSDNSOHNEFEISLRRFRGWIIVLAYTPPNAOHITVLRVY 420
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480

481 QRDITGWKKFVAD-RKKTSGIC 502
474 QLEITTAWLKRVADKXKKTNGVC 496

RESULT 8
3-10-006-852-14
Sequence 14, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
APPLICANT: Kinnerseley, Alan M.
APPLICANT: Turano, Frank J.
TITLE OF INVENTION: Methods for Regulating Plant GABA Production
FILE REFERENCE: 7224-65
CURRENT APPLICATION NUMBER: US/10/006,852
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: US 60/246,367
PRIOR FILING DATE: 2000-11-07
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent in version 3.1
SEQ ID NO 8
LENGTH: 493
TYPE: PRT
ORGANISM: Arabidopsis thaliana
CURRENT APPLICATION NUMBER: US/10/006,852

Query Match 84.1%; Score 2199; DB 15; Length 496;
Best Local Similarity 84.1%; Pred. No. 2e-204;
Matches 423; Conservative 34; Mismatches 38; Indels 8; Gaps 3;

1 MVLHVSSEDSVSHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
1 MVLKSTASESDVSHSTFASRYVTSLSRFRKMPENSIPKEAAYQIINDEMLDGNPRINL 60
61 ASFVTTWMEPECDKLIIMSSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
61 ASFVTTWMEPECNTLMWDSINKNYVDMDEYVTTTELQNRNVNMIHLFNAPLEAEFATVG 120
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
121 VGTGSSRAIMLAGLAFKRKWKQKKAQKGFDPDENIVTGANVQVCWEKFPARYFEVELKE 180
181 VKLSEGYVMDPQQAQVDMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
181 VKLSDGYVMDPEKAVEMVDENTICVAAILGSLNGEPEFVKLNDLLVERKNETGWDTP 240
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
241 IHVDAASGGFTAPFLYPELEWDFRLEPLVKSINVSCHKYGLVYAGIGWVWNRKEDLPEEL 300
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
301 IFHINYLGDQPTFLNFSKGSQVIAQYQLIRLGHGEGYRNWENCRNMIVLRREGLEK 360
361 TERFNIVSKDEGVPLVAFSLKSDSSCHTEFEISDMLRRYGVIPAYTPPNAOHITVLRVY 420
361 SGRFNIIKEIGVPLVAFSLKSDNSOHNEFEISLRRFRGWIIVLAYTPPNAOHITVLRVY 420
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480
421 IREDFSRTLAERLVIDIEKVMRELDLPSRVHKKISLQGEKSESNDMLMVTYVKSIDK 480

481 QRDITGWKKFVAD-RKKTSGIC 502
474 QLEITTAWLKRVADKXKKTNGVC 496

Query Match 81.7%; Score 2135.5; DB 15; Length 493;
 Best Local Similarity 82.4%; Pred. No. 2.9e-198;
 Matches 411; Conservative 32; Mismatches 45; Indels 11; Gaps 2;

1 MVLSHAVSESDVSVHSTFASRYVTSIPRFKMPENSIIPKGAAYQIINDELMLDGNPRLNL 60
 1 MVLSTKTVSESDVSVHSTFASRYVRSNIPREFMPENSIIPKGAAYQIINDELMLDGNPRLNL 60

61 ASFVTTWMEPECDKLMSSINKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEBAETA 120
 61 ASFVTTWMEPECDKLMWESINKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEBAETA 120

121 VGTVGSSEAIMLAGLAFKRWQNKRAEGKPYDKENIVTGANVQVCWEKFAFYFEVLEKE 180
 121 VGTVGSSEAIMLAGLAFKRWQNKRAEGKPYDKENIVTGANVQVCWEKFAFYFEVLEKE 180

181 VKLSEGYVYVMDPQAAVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKNKKEIGW 240
 181 VNLREDIYVMDPVKAVEMVDENTICVAAALGSLTNGEFEDVKLLNDLLVEKNKKEIGW 240

241 IHVDAASGGFTIAPFLYPELEWDFRLLPVKSIIVSGHKYGLVYAGIWMVWRNKEDLPEEL 300
 241 IHVDAASGGFTIAPFLYPELEWDFRLLPVKSIIVSGHKYGLVYAGIWMVWRNKEDLPEEL 300

301 IFHINYLGADQPTTFLNFSKSSQVIAQYQYQIIRLGHGEGYRNWMCNENMIVLREGLEK 360
 301 IFHINYLGADQPTTFLNFSKSSQVIAQYQYQIIRLGHGEGYRNWMCNENMIVLREGLEK 360

361 TERFNIYSKDEGVPLVAFSLKDSCHTEFEISDMRLRRYGVIPAYTMPNAQHIIVLRVV 420
 361 TGRFKIYSKENGVPVAFSLKDSCHTEFEISDMRLRRYGVIPAYTMPNAQHIIVLRVV 420

421 IREDFSRITLAERLVIDIEKVMRELDLPSRVVHKISLQGEKSESNDMLMVTYKKSIDDK 480
 421 IREDFSRITLAERLVADEKVLHELDLPSRVVHKISLQGEKSESNDMLMVTYKKSIDDK 480

481 ORDIIITGWKKFVADRKTKS 499
 471 QREVTAYWKKLL-ETKTKN 488

SULT 7
 -10-006-852-4
 Sequence 4, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
 APPLICANT: Kinnerseley, Alan M.
 APPLICANT: Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 4
 LENGTH: 494
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 3-10-006-852-4

Query Match 81.5%; Score 2130; DB 15; Length 494;
 Best Local Similarity 80.2%; Pred. No. 1e-197;
 Matches 404; Conservative 46; Mismatches 42; Indels 12; Gaps 3;

1 MVLSHAVSESDVSVHSTFASRYVTSIPRFKMPENSIIPKGAAYQIINDELMLDGNPRLNL 60
 1 MVLTKTAT-NDESVCTMFGSRVTRTLPKYEIGENSIPKDAAYQIINDELMLDGNPRLNL 59

61 ASFVTTWMEPECDKLMSSINKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEBAETA 120
 60 ASFVTTWMEPECDKLMSSINKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEBAETA 119

QY 121 VGTVGSSEAIMLAGLAFKRWQNKRAEGKPYDKENIVTGANVQVCWEKFAFYFEVLEKE 180
 DB 120 VETVGSSEAIMLAGLAFKRWQNKRAEGKPYDKENIVTGANVQVCWEKFAFYFEVLEKE 179

QY 181 VKLSEGYVYVMDPQAAVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKNKKEIGW 240
 DB 180 VNLSEGYVYVMDPQAAVDMVDENTICVAAALGSLTNGEFEDVKLLNDLLVEKNKKEIGW 239

QY 241 IHVDAASGGFTIAPFLYPELEWDFRLLPVKSIIVSGHKYGLVYAGIWMVWRNKEDLPEEL 300
 DB 240 IHVDAASGGFTIAPFLYPELEWDFRLLPVKSIIVSGHKYGLVYAGIWMVWRNKEDLPEEL 299

QY 301 IFHINYLGADQPTTFLNFSKSSQVIAQYQYQIIRLGHGEGYRNWMCNENMIVLREGLEK 360
 DB 300 IFHINYLGADQPTTFLNFSKSSQVIAQYQYQIIRLGHGEGYRNWMCNENMIVLREGLEK 359

QY 361 TERFNIYSKDEGVPLVAFSLKDSCHTEFEISDMRLRRYGVIPAYTMPNAQHIIVLRVV 420
 DB 360 TERFNIYSKDEGVPLVAFSLKDSCHTEFEISDMRLRRYGVIPAYTMPADAGHIIVLRVV 419

QY 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVVHKISLQGEKSESNDMLMVTYKKSIDDK 480
 DB 420 IREDFSRITLAERLVADEKVLHELDLPSRVVHKISLQGEKSESNDMLMVTYKKSIDDK 479

QY 481 Q--RDIITGWKKFVADRKTKS 502
 DB 471 EILMEVIVGWRKFKVDRKTKS 494

RESULT 8
 US-10-006-852-6
 ; Sequence 6, Application US/10006852
 ; Publication No. US20030046732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kinnerseley, Alan M.
 ; APPLICANT: Turano, Frank J.
 ; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 ; FILE REFERENCE: 7224-65
 ; CURRENT APPLICATION NUMBER: US/10/006,852
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: US 60/246,367
 ; PRIOR FILING DATE: 2000-11-07
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 500
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-006-852-6

Query Match 79.5%; Score 2080; DB 15; Length 500;
 Best Local Similarity 78.8%; Pred. No. 7.3e-193;
 Matches 395; Conservative 44; Mismatches 52; Indels 10; Gaps 1;

QY 1 MVLSHAVSESDVSVHSTFASRYVTSIPRFKMPENSIIPKGAAYQIINDELMLDGNPRLNL 60
 DB 1 MVLSTKTVSESDVSVHSTFASRYVRSNIPREFMPENSIIPKGAAYQIINDELMLDGNPRLNL 60

QY 61 ASFVTTWMEPECDKLMSSINKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEBAETA 120
 DB 61 ASFVTTWMEPECDKLMWESINKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEBAETA 120

QY 121 VGTVGSSEAIMLAGLAFKRWQNKRAEGKPYDKENIVTGANVQVCWEKFAFYFEVLEKE 180
 DB 121 VGTVGSSEAVMLAGLAFKRWQNKRAEGKPYDKENIVTGANVQVCWEKFAFYFEVLEKE 180

QY 181 VKLSEGYVYVMDPQAAVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKNKKEIGW 240
 DB 181 VKLREGYVYVMDPQAAVDMVDENTICVVAALGSLTNGEFEDVKLLNDLLVEKNKKEIGW 240

QY 241 IHVDAASGGFTIAPFLYPELEWDFRLLPVKSIIVSGHKYGLVYAGIWMVWRNKEDLPEEL 300

241 IHVDAASGGFIAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGWVWWRKTDLPDEL 300
 301 IFHINYLGAQDQPTFLNFKSGSSQVIAQYYQLIRLGHGEGYRNWMCNENMIVIREGLEK 360
 301 IFHINYLGAQDQPTFLNFKSGSSQVIAQYYQLIRLGHGEGYRNWMCNENMIVIREGLEK 360
 361 TERFNIVSKDEGVPVAFSLKDSCHTEFEISDMRLRYRKGWIVPAYTMPNAQHITVLRVV 420
 361 TGRFNIVSKDEGVPVAFSLKDSCHTEFEISDMRLRYRKGWIVPAYTMPNAQHITVLRVV 420
 421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNSDNLMTVTKKSIDK 480
 421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNSDNLMTVTKKSIDK 480
 481 QRDITIGMKKFFVADRKRKTSGI 501
 471 QREVTAYWKKFVDRKTDKNGV 491

Sequence 2, Application US/10167547C
 Publication No. US20030170653A1
 GENERAL INFORMATION:
 APPLICANT: E.I. du Pont de Nemours and Company
 APPLICANT: Damude, Howard G.
 TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
 TITLE OF INVENTION: Butyrolactone and its Intermediates
 FILE REFERENCE: CL1804 US NA
 CURRENT APPLICATION NUMBER: US/10/167,547C
 CURRENT FILING DATE: 2003-03-17
 PRIOR APPLICATION NUMBER: 60/297198
 PRIOR FILING DATE: 2001-06-08
 NUMBER OF SEQ ID NOS: 67
 SOFTWARE: Microsoft Office 07

SEQ ID NO 2
 LENGTH: 498
 TYPE: PRT
 ORGANISM: *Alstroemeria*
 -10-167-547C-2

Query Match 79.3%; Score 2073; DB 12; Length 498;
 Best Local Similarity 77.4%; Pred. No. 3.5e-192;
 Matches 391; Conservative 53; Mismatches 51; Indels 10; Gaps 2;

1 MVLGSHVSEDSVSHSTFASRYVRSLSFRFKMPENSIPKEAAYQIINDELMDGNPRLN 60
 1 MVLSSAVSDNTGPHVCTFASRYVRSLSFRFKMPENSIPKDTAYQIVNDELMDGNPRLN 60
 61 ASFVTTWMEPECDKLIMSS INKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEAEFTAVG 120
 61 ASFVTTWMEPECDKLIMSS INKNYVDMDEYVPTTELQNRQVNMIAHLFNAPIGDEETAVG 120
 121 VGTGSSSAIMLAGLAFKRWKQNRKAEKGPVDPNIVTGANVQVCEKFKARYEVELKE 180
 121 VGTGSSSAIMLAGLAFKRWKQNRKAEKGPVDPNIVTGANVQVCEKFKARYEVELKE 180
 181 VKLSEGYVMPDQAVDMVDENTICVADILGSLNGEFEDVKLLNDLLEKKNKGTGDDTP 240
 181 VKLREGYVMPDQAVDMVDENTICVAAALGSLTGTGFEFVKLLNELLTKKNETGWET 240
 241 IHVDAASGGFIAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGWVWWRKTDLPDEL 300
 241 IHVDAASGGFIAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGWVWWRKTDLPDEL 300
 301 IFHINYLGAQDQPTFLNFKSGSSQVIAQYYQLIRLGHGEGYRNWMCNENMIVIREGLEK 360
 301 IFHINYLGAQDQPTFLNFKSGSSQVIAQYYQLIRLGHGEGYRNWMCNENMIVIREGLEK 360
 361 TERFNIVSKDEGVPVAFSLKDSCHTEFEISDMRLRYRKGWIVPAYTMPNAQHITVLRVV 420
 361 TGRFNIVSKDEGVPVAFSLKDSCHTEFEISDMRLRYRKGWIVPAYTMPNAQHITVLRVV 420

421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNSDNLMTVTKKSIDK 477
 421 IREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNSDNLMTVTKKSIDK 477
 478 IDKORDITIGMKKFFVADRKRKTSGI 502
 474 LEIEKVIITHWQVDMGKTKNKVC 498

Sequence 18, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
 APPLICANT: Kinnersely, Alan M.
 APPLICANT: Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 18
 LENGTH: 502
 TYPE: PRT
 ORGANISM: *Lycopersicon esculentum*
 US-10-006-852-18

Query Match 77.0%; Score 2014; DB 15; Length 502;
 Best Local Similarity 76.0%; Pred. No. 1.9e-186;
 Matches 387; Conservative 53; Mismatches 55; Indels 14; Gaps 5;

1 MVLGSHVSEDSVSHSTFASRYVRSLSFRFKMPENSIPKEAAYQIINDELMDGNPRLN 59
 1 MVLTTISIRDSSESLHCTFASRYVRSLSFRFKMPENSIPKEAAYQIINDELMDGNPRLN 60
 60 LASFVTTWMEPECDKLIMSS INKNYVDMDEYVPTTELQNRQVNMIAHLFNAPLEAEFTAV 119
 61 LASFVTTWMEPECDKLIMSS INKNYVDMDEYVPTTELQNRQVNMIAHLFNAPVGDDEETAV 120
 120 GVTGSSSAIMLAGLAFKRWKQNRKAEKGPVDPNIVTGANVQVCEKFKARYEVELKE 179
 121 GVTGSSSAIMLAGLAFKRWKQNRKAEKGPVDPNIVTGANVQVCEKFKARYEVELKE 180
 180 EVKLSSEGYVMPDQAVDMVDENTICVADILGSLNGEFEDVKLLNDLLEKKNKGTGDDTP 239
 181 EVKLSSEGYVMPDQAVDMVDENTICVAAALGSLTGTGFEFVKLLNELLTKKNETGWET 240
 240 PIHVDAASGGFIAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGWVWWRKTDLPDEL 299
 241 PIHVDAASGGFIAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGWVWWRKTDLPDEL 300
 300 LIHINYLGAQDQPTFLNFKSGSSQVIAQYYQLIRLGHGEGYRNWMCNENMIVIREGLEK 359
 301 LVPHINYLGAQDQPTFLNFKSGSSQVIAQYYQLIRLGHGEGYRNWMCNENMIVIREGLEK 360
 360 KTRFRNIVSKDEGVPVAFSLKDSCHTEFEISDMRLRYRKGWIVPAYTMPNAQHITVLRVV 419
 361 KWGRFPIVSKDQVGVVAFSLKDSCHTEFEISDMRLRYRKGWIVPAYTMPNAQHITVLRVV 420
 420 VIREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNSDNLMTVTKKSIDK 473
 421 VIREDFSRTLAERLVIDIEKVMRELDLPSRVIHKISLGOEKSESNSDNLMTVTKKSIDK 473
 474 KKSIDKQDITIGMKKFFVADRKRKTSGI 502
 481 -----ETQKDIKHWKRKAG--KKTSGVC 502

RESULT 11
 US-10-006-852-10

Sequence 10, Application US/10006852

Publication No. US20030046732A1

GENERAL INFORMATION:

APPLICANT: Kinnersely, Alan M.

APPLICANT: Turano, Frank J.

TITLE OF INVENTION: Methods for Regulating Plant GABA Production

FILE REFERENCE: 7224-65

CURRENT APPLICATION NUMBER: US/10/006,852

CURRENT FILING DATE: 2002-07-01

PRIOR APPLICATION NUMBER: US 60/246,367

PRIOR FILING DATE: 2000-11-07

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patent in version 3.1

SEQ ID NO 10

LENGTH: 494

TYPE: PRN

ORGANISM: Arabidopsis thaliana

-10-006-852-10

Query Match 75.3%; Score 1969; DB 15; Length 494;

Best Local Similarity 75.2%; Pred. No. 4.2e-182;

Matches 373; Conservative 56; Mismatches 61; Indels 6; Gaps 3;

1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKAAAYQIINDELMLDGNPRLNL 60

1 MVLA-TNSDSEHLHSTFASRYVRVAVVPRFKMPDHPKDAAYQVINDELMLDGNPRLNL 59

61 ASFTVTWMEPECDKLLMSSINKNYVDMDEVYVTTTELQNRQVNMIAHLFNAPLEAEATAVG 120

60 ASFTVTWMEPECDKLLMDSVKNYVDMDEVYVTTTELQNRQVNMIAHLFNAPVGDDEAIG 119

121 VGTVGSSEAIMLAGLAFKRKQWKRKAEGKPYDKPNIIVTGANVQVCWEKFARYFEVELKE 180

120 CGTVGSSEAIMLAGLAFKRKQWKRKAQGLPIDKPNIVTGANVQVCWEKFARYFEVELKE 179

181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKKNKGTGWDTP 240

180 VKLSEGYVMDPAKAVEMVDENTICVAAILGSLTNGEFEDVKLLNDLLAEKNAETGNETP 239

241 IHVDAASGFTAFPLYPELEWDFRPLVKSINVSCHKYGLVYAGIYVWIRNKEDLPEEL 300

240 IHVDAASGFTAFPLYPELEWDFRPLVKSINVSCHKYGLVYAGVWVWTKDLDPEEL 299

301 IFHNYLGADQPTFTLNFSGSSQVIAQYQIIRLIGHEGYNMNCRENMIIVLRGLEX 360

300 VFHNYLGADQPTFTLNFSGSSQVIAQYQIIRLIGHEGYNMNCRENMIIVLRGLEX 359

361 TERFNIVSKDGVPLVAFSLKSSCHTEFEISDMLRRYGTWIVPAYTPPNAOHIIVLRVV 420

360 TGFNIVSKDGVPLVAFSLKSSCHTVEFEISDMLRRYGTWIVPAYTPPNAOHIIVLRVV 419

421 IREDFSRRLAERLVIDIEKVMRELDLPSRVHIKISLQEKSESNDMLVTVKKSIDK 480

420 IREDFSRGLADRLIHIQVLKEIGLSRIAHLAA---AAAVSGDDEEVKVKTKMWSL 475

481 ORDITGKWKVADRK 496

476 B-DITKYKRLVHRK 490

RESULT 12

-10-167-547C-4

Sequence 4, Application US/10167547C

Publication No. US20030170653A1

GENERAL INFORMATION:

APPLICANT: E.I. du Pont de Nemours and Company

APPLICANT: Damude, Howard G.

TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma

TITLE OF INVENTION: Butyrolactone and its Intermediates

FILE REFERENCE: CL1804 US NA

CURRENT APPLICATION NUMBER: US/10/167,547C

CURRENT FILING DATE: 2003-03-17

PRIOR APPLICATION NUMBER: 60/297198

PRIOR FILING DATE: 2001-06-08

NUMBER OF SEQ ID NOS: 67

SOFTWARE: Microsoft Office 07

SEQ ID NO 4

LENGTH: 509

TYPE: PRN

ORGANISM: Alstroemeria

US-10-167-547C-4

Query Match 71.3%; Score 1864; DB 12; Length 509;

Best Local Similarity 73.3%; Pred. No. 6.9e-172;

Matches 351; Conservative 54; Mismatches 70; Indels 4; Gaps 2;

QY 1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKAAAYQIINDELMLDGNPRLNL 60

Db 1 MALSSVVSDSNNOVQCTYASRYVRDEAPGFRMPKSIKPEAAAFMINDELMLDGNPRLNL 60

QY 61 ASFTVTWMEPECDKLLMSSINKNYVDMDEVYVTTTELQNRQVNMIAHLFNAPLEAEATAVG 120

Db 61 ASFTVTWMEPECDRLMMSTINKNYALMDDYPVTIDIQNRQVNMIAHLFNAPLEAEATAVG 120

QY 121 VGTVGSSEAIMLAGLAFKRKQWKRKAEGKPYDKPNIIVTGANVQVCWEKFARYFEVELKE 180

Db 121 CAIVGSSSEAMLAGLAFKRKQWKRKAEGKPYDKPNIIVTGANVQVCWEKFARYFEVELKE 180

QY 181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKKNKGTGWDTP 240

Db 181 VNLREGYYVMDPEKAVEMVDENTICVAAILGSLTNGEFEDVKLLNDLLVEKKNKGTGWDTP 240

QY 241 IHVDAASGFTAFPLYPELEWDFRPLVKSINVSCHKYGLVYAGIYVWIRNKEDLPEEL 300

Db 241 IHVDAAGIIFAPLYPELEWDFRPLVKSINVSCHKYGLVYAGVWVWIRNKEDLPEEL 300

QY 301 IFHNYLGADQPTFTLNFSGSSQVIAQYQIIRLIGHEGYNMNCRENMIIVLRGLEX 360

Db 301 IFHNYLGIDQPTFTLNFSGSNQIIGQYVYQIIRLIGHEGYNMNCRENMIIVLRGLEX 360

QY 361 TERFNIVSKDGVPLVAFSLKSSCHTEFEISDMLRRYGTWIVPAYTPPNAOHIIVLRVV 420

Db 361 MGVEIISKDIGAPLVIALKDSKSHSVFKIADTIRFGWTTIPAYTPMPKDVBEHIAVLRVV 420

QY 421 IREDFSRRLAERLVIDIEKVMRELDLPSRVHIKISLQEKSESNDMLVTVKKSIDK 479

Db 421 IREDFRSRLAERLANDMKKVLVELDHPHSRT---TTIAHVKAVENTGNVYV-IKKSIVE 475

RESULT 13

US-10-167-547C-6

Sequence 6, Application US/10167547C

Publication No. US20030170653A1

GENERAL INFORMATION:

APPLICANT: E.I. du Pont de Nemours and Company

APPLICANT: Damude, Howard G.

TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma

TITLE OF INVENTION: Butyrolactone and its Intermediates

FILE REFERENCE: CL1804 US NA

CURRENT APPLICATION NUMBER: US/10/167,547C

CURRENT FILING DATE: 2003-03-17

PRIOR APPLICATION NUMBER: 60/297198

PRIOR FILING DATE: 2001-06-08

NUMBER OF SEQ ID NOS: 67

SOFTWARE: Microsoft Office 07

SEQ ID NO 6

LENGTH: 529

TYPE: PRN

ORGANISM: Alstroemeria

US-10-167-547C-6

Query Match 71.3%; Score 1864; DB 12; Length 529;

Best Local Similarity 73.3%; Pred. No. 7.3e-172;

Matches 351; Conservative 54; Mismatches 70; Indels 4; Gaps 2;

QY 1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKAAAYQIINDELMLDGNPRLNL 60

```

1 MALSSVSDSNVQCYASRYRDEAPGRNPEKSIKBAEATMINDLMDGNPLML 60
61 ASFVTTWMEPCDKLIMSSINKNVDMDEYPTTELQNRQVNMIAHLFNAPLEAEATVAG 120
61 ASFVTTWMEPCDKLIMSSINKNVDMDEYPTTELQNRQVNMIAHLFNAPLEAEATVAG 120
121 VGTVGSSEIAMLGLAFKRWKQKRAEGKVPDKPNIVTGANVQVCEKFAFYFVELKE 180
121 CATVGSSEAMMLAGLAFKRWKQKRAEGKPYDKENMVTGSNVQVWVFAKPYFVEMKK 180
181 VKLSEGYVMDPQAVDMVDENTICVADIIIGSTLNGEFEDYVKLINDLIVKKNKTKGMDTP 240
181 VNLREGYVMDPEKAVEMVDENTICVAAIIGSTLTGFEFEDVVKLINDLIVKKNKTKGMDTP 240
241 IHVDAASGGFIAPLYPELSDWDFRLPLVKSINVSCHKYGLVYAGIGWIVRNKEDLPEEL 300
241 IHVDAALGGFIAPLYPELSDWDFRLPLVKSINVSCHKYGLVYAGIGWIVRNKEDLPEEL 300
301 IFHNYLGAQDPTFLNFKSGSQVIAQYQIIRLGLGHEGYNMNCRENMIVLRGLEK 360
301 IFHNYLGIQDPTFLNFKSGSQVIAQYQIIRLGLGHEGYNMNCRENMIVLRGLEK 360
361 TERNIVSKBEGVPLVAFSLKDSCHTEPELSDMLRYYGIVPAYTTPNAQHITVLRV 420
361 MGVEEITSKDIGAPLVIALKDSKSHVFKIADTIRRFPGWTIPAYTMPKDVHIAVLRV 420
421 IREDFSRDLAERLVIDIEKVMRELDLPSRVVIKHSISLQKESNSDNLMTVVKSDID 479
421 IREDFSRDLAERLVIDIEKVMRELDLPSRVVIKHSISLQKESNSDNLMTVVKSDID 479

```

```

SUIT 14
-10-156-761-12189
Sequence 12189, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 12189
LENGTH: 470
TYPE: PRN
ORGANISM: Streptomyces avermitilis
-10-156-761-12189

```

```

Query Match 41.4%; Score 1082; DB 15; Length 470;
Best Local Similarity 49.8%; Pred. No. 4.7e-96;
Matches 209; Conservative 74; Mismatches 131; Indels 6; Gaps 4;
28 PRFQKPNISPKAAAYQIINDELMDGNPRNLASFTVWMEPCDKLIMSINKNYVDM 87
35 PGRHLDGGLPSTAYQLVHDELMDGNARLNLATFTVWMEPCQAGLVMAEQDKNMDK 94
88 DEYPTTELQNRQVNMIAHLFNAPLEAEATVGTVGSSEAMLAGLAFKRWKQKRA 147
95 DEYPTTAELEKRCVAMLADLWNAF--DAGAAVGCSTTGSSEACMLAGLAKRWA--RRNA 151
148 EGKPYD--KPNIVTGANVQVCEKFAFYFVELKEVLSGYSVMDPQAVDMVDENTIC 205

```

```

Db 152 DRYPARDVNPNLVMGINVQVCEKFCNFWEMRQVPLEGERYHLDPGAAELCDENTIG 211
QY 206 VADIILGSLNGFEFVVKLINDLIVKKNKTKGMDTPIHVDAASGGFIAPPLYPELEWDFRL 265
Db 212 VVGIILGSLNGFEFVVKLINDLIVKKNKTKGMDTPIHVDAASGGFIAPPLYPELEWDFRL 271
QY 266 PLVKINVSCHKYGLVYAGIGWIVRNKEDLPEELIPIHNYLGAQDPTFLNFKSGSQV 325
Db 272 PRVAVINTSGHKYGLVYAGIGWIVRNKEDLPEELVFRVNYLGGNMPFALNFRPQAV 331
QY 326 IAQYQIIRLGLGHEGYNMNCRENMIVLRGLEKTERFNIYSKDEGVPLVAF--SLKDS 384
Db 332 VAQYTFELGREGFRVQOITRNVARSLAERVAALGDFHLLTRGDELFPVFAFTAPVA 391
QY 385 CHTEPEISDMLRYYGIVPAYTTPNAQHITVLRVVIREDFSRDLAERLVIDIEKVMREL 444
Db 392 SYDVFVDSRRMEHGLWVPAFTPPNRELDLVLVCRNGFSTDLAELFVEDLSRLPLDL 451

```

```

RESULT 15
US-10-156-761-11138
Sequence 11138, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11138
LENGTH: 454
TYPE: PRN
ORGANISM: Streptomyces avermitilis
US-10-156-761-11138

```

```

Query Match 40.2%; Score 1051; DB 15; Length 454;
Best Local Similarity 45.6%; Pred. No. 4.5e-93;
Matches 198; Conservative 86; Mismatches 144; Indels 6; Gaps 3;
QY 14 VHSFTASRYVTSLSRPFKMPENSIPKAAAYQIINDELMDGNPRNLASFTVWMEPCD 73
Db 6 VAALFGNRFLETPAPSQTFPEEGMTATDTRMLLDELVMEGDFQRLNLAFTVWMEPEAQ 65
QY 74 KLTMSSINKNVDMDEYPTTELQNRQVNMIAHLFNAPLEAEATVGTVGSSEAMLA 133
Db 66 RIIAEMLHRNFIDHAEYFISAEIEQRCVRLADLFRAP---GKTTGCRTOGSSSEAIMLG 121
QY 134 GLAFKRWKQKRAEGKVPDKPNIVTGANVQVCEKFAFYFVELKEVLSGYSVMDPQ 193
Db 122 ALSLKWKRERRQAAANLPADRPNLVFGDGHVWVEKFCRYFDVPEPRIVPLAEADKYTIGPE 181
QY 194 QAVDMVDENTICVADIIIGSTLNGEFEDVVKLINDLIVKKNKTKGMDTPIHVDAASGGFIAP 253
Db 182 DVEPHIDENTIGVAVVGTFTGHKDDVVGDIDKLLRDRVTKERDLDFIPIHVDGASGGFWP 241
QY 254 FLYPELEWDFRLPLVKSINVSCHKYGLVYAGIGWIVRNKEDLPEELIPIHNYLGAQDPT 313
Db 242 FLYPDSKWDPRLEQVRSINVSCHKYGLVYAGIGWIVRNKEDLPEELIPIHNYLGAQDPT 301
QY 314 FTLNFKSGSQVIAQYQIIRLGLGHEGYNMNCRENMIVLRGLEKTERFNIYSKDEGVPLVAF--EG 372
Db 302 FTLNFKSGSQVIAQYQIIRLGLGHEGYNMNCRENMIVLRGLEKTERFNIYSKDEGVPLVAF--EG 372

```


GenCore version 5.1.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.
 M protein - protein search, using sw model
 un on: October 22, 2003, 15:04:24 ; Search time 42 seconds
 (without alignments)
 505.715 Million cell updates/sec
 title: US-10-006-852-2
 effect score: 2615
 sequence: 1 MVLSHAVSSDVSVHSTFAS.....DIITGKKFVADRKKTSGIC 502
 coring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 searched: 328717 seqs, 42310858 residues

total number of hits satisfying chosen parameters: 328717

minimum DB seq length: 0
 maximum DB seq length: 2000000000
 post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

database : Issued Patents AA: *
 1: /cgn2_6/prodata/2/iaa/5A COMB.pep: *
 2: /cgn2_6/prodata/2/iaa/5B COMB.pep: *
 3: /cgn2_6/prodata/2/iaa/6A COMB.pep: *
 4: /cgn2_6/prodata/2/iaa/6B COMB.pep: *
 5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep: *
 6: /cgn2_6/prodata/2/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1023	39.1	466	3	US-09-068-195-25
2	325	12.4	589	4	US-08-939-309-8
3	325	12.4	589	4	US-09-849-180-8
4	325	12.4	589	4	US-09-356-643B-2
5	288	11.0	552	4	US-09-356-643B-11
6	264.5	10.1	542	4	US-08-939-309-6
7	264.5	10.1	542	4	US-09-849-180-6
8	264.5	10.1	542	4	US-09-356-643B-4
9	246.5	9.4	568	4	US-08-939-309-2
10	246.5	9.4	568	4	US-09-849-180-2
11	246.5	9.4	568	4	US-09-356-643B-6
12	227	8.7	568	4	US-08-939-309-4
13	227	8.7	568	4	US-09-849-180-4
14	227	8.7	568	4	US-09-356-643B-8
15	224	8.6	568	3	US-09-238-373-2
16	224	8.6	568	4	US-09-740-369-2
17	160.5	6.1	488	4	US-08-939-309-10
18	160.5	6.1	488	4	US-09-849-180-10
19	160.5	6.1	488	4	US-09-356-643B-10
20	133.5	5.1	525	4	US-09-328-352-4492
21	116	4.4	509	3	US-09-377-557-20
22	107.5	4.1	662	3	US-08-779-814-5
23	105	4.0	729	4	US-09-134-001C-4728
24	104.5	4.0	784	4	US-09-004-838-12
25	101.5	3.9	1221	4	US-09-107-532A-3959
26	101	3.9	730	3	US-09-398-865A-2
27	101	3.9	730	4	US-09-710-714-2

28	100.5	3.8	398	4	US-09-328-352-4507	Sequence 4507, Ap
29	100.5	3.8	544	3	US-08-687-590-30	Sequence 30, Appl
30	99.5	3.8	393	3	US-09-377-557-14	Sequence 14, Appl
31	99.5	3.8	458	4	US-09-134-001C-4422	Sequence 4422, Ap
32	99.5	3.8	529	3	US-09-464-483-2	Sequence 2, Appl
33	99.5	3.8	529	3	US-09-414-664-2	Sequence 2, Appl
34	99	3.8	419	4	US-09-328-352-7844	Sequence 7844, Ap
35	98.5	3.8	1447	4	US-09-376-330-17	Sequence 17, Appl
36	98	3.7	623	4	US-09-107-532A-4726	Sequence 4726, Ap
37	97.5	3.7	402	3	US-09-464-483-4	Sequence 4, Appl
38	97.5	3.7	402	3	US-09-414-664-4	Sequence 4, Appl
39	97	3.7	428	3	US-09-052-778-12	Sequence 12, Appl
40	97	3.7	428	4	US-09-134-001C-4879	Sequence 4879, Ap
41	97	3.7	990	4	US-09-627-376-7	Sequence 7, Appl
42	97	3.7	1120	4	US-09-147-404-1	Sequence 1, Appl
43	96.5	3.7	597	4	US-09-252-991A-32073	Sequence 32073, A
44	95.5	3.7	391	3	US-09-377-557-18	Sequence 18, Appl
45	95.5	3.7	508	4	US-09-858-664A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
 ; Sequence 25, Application US/09068195B
 ; Patent No. 6140078
 ; GENERAL INFORMATION:
 ; APPLICANT: Sanders, Jan W.
 ; APPLICANT: Ledesboer, Adrianus M.
 ; APPLICANT: Venema, Gerard
 ; APPLICANT: Kok, Jan
 ; TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid
 ; TITLE OF INVENTION: Bacterium, and Its Use in a Lactic Acid Bacterium for
 ; TITLE OF INVENTION: Production of Desired Protein
 ; FILE REFERENCE: Sanders-60113/0252227
 ; CURRENT APPLICATION NUMBER: US/09/068,195B
 ; EARLIER FILING DATE: 1998-07-29
 ; EARLIER APPLICATION NUMBER: PCT/EP97/04755
 ; EARLIER FILING DATE: 1997-08-20
 ; EARLIER APPLICATION NUMBER: EP 97200744/7
 ; EARLIER FILING DATE: 1997-03-13
 ; EARLIER APPLICATION NUMBER: EP 96202444/4
 ; EARLIER FILING DATE: 1996-09-05
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 25
 ; LENGTH: 466
 ; TYPE: PRT
 ; ORGANISM: L. lactis MGI363
 US-09-068-195-25

Query	Match	Score	DB 3;	Length	466;
18	FASRYVTSIPRFKMPENSIPKAAQIINDELMLDGNPFLMLASVTTWMEPECDKLM 77	39.1*			
18	FGSEQVDLPKYKLAQOQIEPRVAYQLVODEMLDEGNARLNLAATFCQTYMPEFAVKLMS 77	45.0*			
78	SSINQNVDMDEYPTTTELQNRQVNIAMHFNAPLEAEATVAVGVGVTSSEATMLAGLAP 137				
78	QTEKNAIDKSEYPRITEENRCVNNIADLWNA--SEKGIYGTSTIGSSEACMLGGMMAN 135				
138	KRWQNRKRAEGKFDV--KPNIVGTANVQVCKEKFARYFEVELKVKLSEGYVMDPQQA 195				
136	KFSWRKRAEKGLGIDINAKKPNLVISGYQVCKEFCVYWDIEMREVPMDREHMSINLEKV 195				
196	VDVYDENTICVADILGSLNGEFEDVKLNDLLVEKNKTEGWDTPHVDAAAGGFAPFL 255				
196	MDYVDEYTIQVGMINGITTYGRYDDIKALDNLIIEYKQTDYKVIHVDAASGGLYAPFV 255				
256	YPELEMDFRPLVKSINVSCHKYGLVAVAGVIGWVWRNKEDLPBELIFHINYLGAQPPTFT 315				

256 EPELEWDFRLKXNVI... 315
 316 LNFSGSSVIAQYQLRL... 375
 316 INFHSASOLGQYFV... 375
 376 VAFSLKDSCH--TE... 433
 376 VCYKLENSRGNW... 433
 434 VIDIEKVMPELDE... 446
 436 VQDMQRAIDALNK... 448

RESULT 2
 3-08-939-309-8
 Sequence 8, Application US/08939309
 Patent No. 6423527
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 APPLICANT: Zhou, Jianhui
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 TITLE OF INVENTION: POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 METHODS OF USE THEREFOR
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/939,309
 FILING DATE: 29-SEP-1997
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: David, Maki J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 200116.402
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 589 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

-08-939-309-8
 Query Match 12.4%; Score 325; DB 4; Length 589;
 Best Local Similarity 24.7%; Pred. No. 6.2e-24;
 Matches 122; Conservative 98; Mismatches 188; Indels 86; Gaps 20;
 3 LSHAVSESDVSVHSTFASRVR--TSLPRF-KMPENSIFKEAAQYIINDELMDGNPRLN 59
 106 LRGTVEKVKQKQSIQDELRISDSQAMFPQLPSNGIPQDDVIEELN-----KLN 156
 60 LASFVITWMEPEC-----DKLIMSSIN-----KNYVDWDEYVPTTELQNRVMMI 104
 157 DLIPHCQKQKGVSGAVYHGDDLLIHLQIYAYEKYCVANQLHPDVPFPAVKMESEVSVW 216
 105 AHLFNAPLEBAFTAVGVVGSSEAIMLAGLAFKRWQNKRAKAGKVPDKPVIIVGANVQ 164
 217 LRWFNAP---SDTCGCTTSSGGTESLILLACLAKRMVALHHR-----GITEPEIIAFVTAH 268

QY 165 VCWEKFAFYFEVLEKVKLSEGYVYMDPOQAVDMVDENTICVADILGSLN---GEFBDV 221
 DB 269 AGFDKAAAYFGKLRHVELDPTTYQVDLQKVKFKINKTIL---LVGSAPFFPHGIADDI 325
 QY 222 KLNDLLVLEKVKETGMDPTPIHVDAASGGFIAPFL---YPELE-WDFRLPLVKSINVSGH 276
 DB 326 EGLGK-IAQKYK----LPLHVDSCIGSFIVSFMEKAGYKMLPLLDLFRVPGVTSICDTH 379
 QY 277 KYGLVYAGIGWIRNKEDLPEELIFHINYLGDQPTFLNFKSGSSOVIQYQYQLIBLG 336
 DB 380 KYGFAPKSSVIMYRNSDURMHQYVYVPAWTGELGSPFLAGSRECALVVGCVATWNVWG 439
 QY 337 HEGYRNVNENRENMIVLREGLEKTERFNI----VSKDEGVPLVAFSLKDSCHTEFEFIS 392
 DB 440 ENGY---IESCOE-IVGAAMKFKYIQENIPDLNIMGNPFRYSVIFSFKTLNIH---ELS 492
 QY 393 DMLRRYGNIVPAYTTPNAQHITVLRVIREDFSLAERLVIDIEKVMRELDLPSRVI 452
 DB 493 DRLSKKGWHFNALQKP-----VALHMAFTR-LSAHVYVEICDILR----- 531
 QY 453 HKISLGOEKSESNS 466
 DB 532 --TVQVELKSESNS 543

RESULT 3
 US-09-849-180-8
 Sequence 8, Application US/09849180
 Patent No. 6495359
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 APPLICANT: Zhou, Jianhui
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 METHODS OF USE THEREFOR
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seed Intellectual Property Law Group
 STREET: 701 Fifth Avenue, Suite 6300
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98055

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/849,180
 FILING DATE: 04-May-2001
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Pepe, Jeffrey C.
 REGISTRATION NUMBER: 46,985
 REFERENCE/DOCKET NUMBER: 200116.402
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4500
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 589 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-09-849-180-8

Query Match 12.4%; Score 325; DB 4; Length 589;
 Best Local Similarity 24.7%; Pred. No. 6.2e-24;
 Matches 122; Conservative 98; Mismatches 188; Indels 86; Gaps 20;

3 LSHAVSESDVSHSTFASRYR--TSLPRF-KMPENSTPKAAAYQIINDELMLDGNRLN 59
 106 LRGTEVEKVKQSIEDLIRSDQSMNFPQLPSNGIPDDVIEELN-----KLN 156
 60 LASFVTTWMEPEC-----DKLIMSSIN-----KNYVDMDEYVTTTELQNRVNM 104
 157 DLIPIHTQWKEGVSAVYHGGDDLHLQTIAYEKYCVANQLHPDVPFAVRKWESEVSMV 216
 105 AHLFNAPLEEAFTAVGVGVSSEAIMLAGLAFKRKQNRKAEGKPKVDRPNIVTGANVQ 164
 217 LRMFNAP--SDTGGCTTSGTESLLACLASAKWYALHHR-----GITEPEIIAPVTAH 268
 165 VWKEFARYFEVELKEVKLSEGYVMDPQAVDMVDENTICVADILGSTLN---GEFEDV 221
 269 AGFDKAAAYFGMKLRHVELDPTTYQVDLKGKVKFKINKNTIL---LVGSAPNPHGIADDI 325
 222 KLLNDLLVEKKNKGTWDTPIHVDAASGGFIAPFL---YPELE-WDFRPLPVKSNVSGH 276
 326 EGLGK-IAQKYK-----LPLHVDSCLSGFSIVSFKAGYKNLPLDPRVPGVTSISCDTH 379
 277 KYGLVYAGIGWIMRNKEDLPEELIFHINYLGDQPTFLNFKSGSSOVIAYQYQLIRLG 336
 380 KYGFAPKSSVIMYRNSDLRHOYVNPWPAWTGGLYGSPTLAGSRPGAIVVGCWATWVMG 439
 337 HEGYRNVMNCRENMIVLREGLEKTERENI---VSKDEGVPIVAFSLKDSCHTEFEIS 392
 440 ENGY---IESCOE-IVGAAMFKKYIQENIPDLNMGPNRYSVIFSFKTLNHH---ELS 492
 393 DMLRRYGVIPAYTPPPNAQHITLVRVIREDFSRTLAERLVIDIEKVMRELDLPSRVI 452
 493 DRLSKKGWFHNAQKPE-----VALHMAFTR-LSAHVVDEICDILR-----531
 453 HKISLQEKSESNS 466
 532 --TTVQELKSESNS 543

SULT 4
 -09-356-643B-2
 Sequence 2, Application US/09356643B
 Patent No. 6569666
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 TITLE OF INVENTION: SHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
 TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
 TITLE OF INVENTION: METHODS OF USE THEREFOR
 FILE REFERENCE: 200116.402C1
 CURRENT APPLICATION NUMBER: US/09/356.643B
 CURRENT FILING DATE: 1999-07-19
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 589
 TYPE: PRT
 ORGANISM: S. cerevisiae
 -09-356-643B-2

Query Match 12.4%; Score 325; DB 4; Length 589;
 Best Local Similarity 24.7%; Pred. No. 6.2e-24;
 Matches 122; Conservative 98; Mismatches 188; Indels 86; Gaps 20;
 3 LSHAVSESDVSHSTFASRYR--TSLPRF-KMPENSTPKAAAYQIINDELMLDGNRLN 59
 106 LRGTEVEKVKQSIEDLIRSDQSMNFPQLPSNGIPDDVIEELN-----KLN 156
 60 LASFVTTWMEPEC-----DKLIMSSIN-----KNYVDMDEYVTTTELQNRVNM 104
 157 DLIPIHTQWKEGVSAVYHGGDDLHLQTIAYEKYCVANQLHPDVPFAVRKWESEVSMV 216
 105 AHLFNAPLEEAFTAVGVGVSSEAIMLAGLAFKRKQNRKAEGKPKVDRPNIVTGANVQ 164
 217 LRMFNAP--SDTGGCTTSGTESLLACLASAKWYALHHR-----GITEPEIIAPVTAH 268

QY 165 VWKEFARYFEVELKEVKLSEGYVMDPQAVDMVDENTICVADILGSTLN---GEFEDV 221
 Db 269 AGFDKAAAYFGMKLRHVELDPTTYQVDLKGKVKFKINKNTIL---LVGSAPNPHGIADDI 325
 QY 222 KLLNDLLVEKKNKGTWDTPIHVDAASGGFIAPFL---YPELE-WDFRPLPVKSNVSGH 276
 Db 326 EGLGK-IAQKYK-----LPLHVDSCLSGFSIVSFKAGYKNLPLDPRVPGVTSISCDTH 379
 QY 277 KYGLVYAGIGWIMRNKEDLPEELIFHINYLGDQPTFLNFKSGSSOVIAYQYQLIRLG 336
 Db 380 KYGFAPKSSVIMYRNSDLRHOYVNPWPAWTGGLYGSPTLAGSRPGAIVVGCWATWVMG 439
 QY 337 HEGYRNVMNCRENMIVLREGLEKTERENI---VSKDEGVPIVAFSLKDSCHTEFEIS 392
 Db 440 ENGY---IESCOE-IVGAAMFKKYIQENIPDLNMGPNRYSVIFSFKTLNHH---ELS 492
 QY 393 DMLRRYGVIPAYTPPPNAQHITLVRVIREDFSRTLAERLVIDIEKVMRELDLPSRVI 452
 Db 493 DRLSKKGWFHNAQKPE-----VALHMAFTR-LSAHVVDEICDILR-----531
 QY 453 HKISLQEKSESNS 466
 Db 532 --TTVQELKSESNS 543

RESULT 5
 US-09-356-643B-11
 ; Sequence 11, Application US/09356643B
 ; Patent No. 6569666
 ; GENERAL INFORMATION:
 ; APPLICANT: Saba, Julie D.
 ; TITLE OF INVENTION: SHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
 ; TITLE OF INVENTION: METHODS OF USE THEREFOR
 ; FILE REFERENCE: 200116.402C1
 ; CURRENT APPLICATION NUMBER: US/09/356.643B
 ; CURRENT FILING DATE: 1999-07-19
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 552
 ; TYPE: PRT
 ; ORGANISM: C. elegans
 ; US-09-356-643B-11

Query Match 11.0%; Score 288; DB 4; Length 552;
 Best Local Similarity 24.2%; Pred. No. 3.1e-20;
 Matches 119; Conservative 89; Mismatches 208; Indels 76; Gaps 21;
 QY 13 SVHSTFASRYVTSIPLPRKMPENSI-----PKAAAYQIINDELMLDGNRLNLAASFVTTWM 68
 Db 95 SIHKDDKQFISTLPIAFLSQDSIMELAKKYEDYNTFN-----IDGG---RVSGAVYTR 147
 QY 69 EPECDKLIMSSINK---NYVDMDEYVTTTELQNRVNMIAHLFNAPLEEAFTAVGVGTV 124
 Db 148 HAHINLLGKIYEKVAFPNPLPDPFPFGARKKAEELIIMVNLVNGP-----EDSSGSVTS 203
 QY 125 GSSEAIMLAGLAFKRKQNRKAEGKPKVDRPNIVTGANVQVCEKPFARYFEVELKEVKLS 184
 Db 204 GGTESIIMACFSY-----RNRAHSLGIEHPVILACKTAHAFAAKAAHLICGMLRHRPVD 257
 QY 185 EGYVMDPQAVDMVDENTICVADILGSTLN---GEFEDVCLLNDLLVEKKNKGTWDTPI 241
 Db 258 SDRNV-DLKEBERLIDSN-VCV-LVGSAPNFPSTIDPIPEIAKL-----GKKG--IPV 307
 QY 242 HVDAASGGFIAPFL---YPELEWDFRPLPVKSNVSGHKYGLVYAGIGWIMRNKEDLP 297
 Db 308 HVDAALGGFMIPMDAGYLIIVDFRFPVTSISCDTHKYGCTPKGSSIVMRSKE---364
 QY 298 EELIFHINYL-----GADQPTFLNFKSGSSOVIAYQYQLIRLGHEGYRNVMNCRENM 351
 Db 365 --LHHFQVFSVADVWCGGIYATPIAGSRAGANTAVAWATLLSFRDEVRRCAQIVKHT 421

352 IVLREGLKTERNIVSKDEGVPLVAFSLKDSCHTEFEI SDMLRRYGVWIPAYTMPNA 411
 422 RMLAEKIEKIKKPYGKSD-VSLVAFSGNGVNI---YEVSDKMKLGNWLNLT-LQNPAA 476
 412 QHITLVRVIREDFSRFLAERLVIDIEKVMRELDLPSRVHVKISLGOEKSESNSD--N 468
 477 IHICLTINQANEVWNAFA---VDLEKICEEL-----AAKGBQKADSGMAAMYG 522
 469 LMVTVKKSIDK 480
 523 MAAQVPKSVVDE 534

RESULT 6
 5-08-939-309-6
 Sequence 6, Application US/08939309
 Patent No. 6423527
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 APPLICANT: Zhou, Jianhui
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 TITLE OF INVENTION: POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 TITLE OF INVENTION: METHODS OF USE THEREFOR
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED AND BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/939,309
 FILING DATE: 29-SEP-1997
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: David, Maki J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 200116.402
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 542 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 -08-939-309-6

Query Match 10.1%; Score 264.5; DB 4; Length 542;
 Best Local Similarity 22.3%; Pred. NO. 7.2e-18;
 Matches 109; Conservative 93; Mismatches 207; Indels 79; Gaps 20;
 3 LSHAVSESDVSVHSTFASRYVTSLSRFRKPKENSIPKEAA-YQIINDELMLDGNPRLNLA 61
 77 LNEVKDELEKSLRIVDRSTEVFTTIPSHSVGRTEVLRALAAIYDDLGGPAPFLGG--RVSGA 134
 62 SF-----VTTWMEPECDKLIMSSINKYVVDVDEFPVTEQLONRCVNNMI 104
 135 VFNREDDKDEREMVEYVFGKFAWTLNLPKLI-----PPGVRIWEAEVVRMC 180
 105 AHLFNAPLEAEATVAVGTVGSGSEAIMLAGLAFKRWKQNKRAEGKRFVDPKNIIVTGANVQ 164
 181 CNMANG--DSEIC-GTWSGTGGISILLACLAR---NELLKRGKYTE--WIVPSSVH 230
 165 VCWEKFARYEVELKVEKLSGYYVMDPQAVMDVENTICVADILGLSTLN---GEFFEDV 221

Db 231 AAFFKAAECFRIVKVKI PVDVPTFKVDLVKMKAAINKRT-CM--LVGSAENPFGTVDDI 287
 Qy 222 KLNDLLVKNKTEGWDPIHVDAAASGFIAPFL-YPELEWDFRPLVYKVINVSHKVKGL 280
 Db 288 EAIGQLGLF-----YDIPVHVDACLGQGLLFFLEEDIRYDFRVPVGVSSISADSHKVKGL 341
 Qy 281 VVAGIGWVWRNKEDLPBELIFHINVLGADQPTFLNFSKGSQVIAQYQIILRIGHEGY 340
 Db 342 APKSSVLYRNKELLENQYFCDADWQGGIYASATMEGSRAGHNIALCWAAMLYHAQSGY 401
 Qy 341 ----RNVMENRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSCHTEFEISDMLR 396
 Db 402 KAMARKIVDTTRK---IRNGLSNIKGKLGQSPD-VCIIVSWTINDGV--ELYRPHNPMK 454
 Qy 397 RYGVWIPAYTMPNNAOHTVLRVIREDFSRFLAERLVIDIEKVMRELDLPSRVHVKIS 456
 Db 455 EKHWQLNGLOFFAGV-HI----VVTWHTHFGLAFAFVADCRAAVEFVKS-----HKPS 503

Qy 457 LGOEKSES 464
 Db 504 ESDKTSEA 511

RESULT 7
 US-09-849-180-6
 Sequence 6, Application US/09849180
 Patent No. 6495359
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 APPLICANT: Zhou, Jianhui
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 METHODS OF USE THEREFOR
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seed Intellectual Property Law Group
 STREET: 701 Fifth Avenue, Suite 6300
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98055

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/849,180
 FILING DATE: 04-May-2001
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Pepe, Jeffrey C.
 REGISTRATION NUMBER: 46,995
 REFERENCE/DOCKET NUMBER: 200116.402
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 542 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 US-09-849-180-6

Query Match 10.1%; Score 264.5; DB 4; Length 542;
 Best Local Similarity 22.3%; Pred. NO. 7.2e-18;
 Matches 109; Conservative 93; Mismatches 207; Indels 79; Gaps 20;
 3 LSHAVSESDVSVHSTFASRYVTSLSRFRKPKENSIPKEAA-YQIINDELMLDGNPRLNLA 61
 77 LNEVKDELEKSLRIVDRSTEVFTTIPSHSVGRTEVLRALAAIYDDLGGPAPFLGG--RVSGA 134

62 SF-----VTTWMEPECDKLMSSINKNYVDMDBYPTTELQRCVYMI 104
 135 VFNREDDKDEREMEEVFGKFAWTPLWPKL-----FPQVRIWEAEVVRMC 180
 105 AHLFNAPLEAEATAVGVTVGSSEAIMLAGLAFKRWKQKKAEGKPKVDKPNIVGTANVQ 164
 181 CNMNG--DSETC-GTWTGSGISILLACLAR-----NRLKRGKYTE--MIVPSSVH 230
 165 VCWEKFAFYFVELKEVKLSBEGYVMDPQAQVDMVDENTICVADILGSLN--GEFEDV 221
 231 AAFKAAECFRKVRKIPVDFVTFKVDLVKMAAINKRT-CM--LVGSAPNFPFGTVDDI 287
 222 KLLNDLLVEKNKETGWDTPHVDAAAGGFIAPFL-YPELEWDFRLPLVKSINVSCHKYGL 280
 288 EAIGQLGLE-----YDIPVHVDAICLGGFLPFLEDEIRYDFRVPGVSSISADSHKYL 341
 281 VYAGIGWVIRWKNEDLPBELIHPHINVLGADQPTFLNFSKSSQVIAQYYQLIRLGHGY 340
 342 APKGSVLYRNKELLHNQYFCDAWQGGIYASATMEGSRAGHNIALCWAAMLXHAQEGY 401
 341 ----RNVMNCRENNMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSCHTEFEISDMLR 396
 402 KANARKIVDTRK----IRNGLSNIKIGIKLOGPSP--VCIVSWTTNDGV--ELYRPHFMK 454
 397 RYGVWVPAVMPNQAQHTVLRVWIREDFSRTLAERLVIDIEKVMRELDLPSRVHKIS 456
 455 EKHQWLNGLQFPAGV-HI-----MVTMNHHTHPLGLAEAFVADCRAAVEFVKS-----HKPS 503
 457 LGQEKSES 464
 504 ESDKTSEA 511

SULT 8
 -09-356-643B-4
 Sequence 4, Application US/09356643B
 Patent No. 6569666
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
 TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
 TITLE OF INVENTION: METHODS OF USE THEREFOR
 FILE REFERENCE: 200116.402C1
 CURRENT APPLICATION NUMBER: US/09/356,643B
 CURRENT FILING DATE: 1999-07-19
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 4
 LENGTH: 542
 TYPE: PRI
 ORGANISM: C. elegans
 -09-356-643B-4
 Query Match 10.1%; Score 264.5; DB 4; Length 542;
 Best Local Similarity 22.3%; Pred. No. 7,2e-18;
 Matches 109; Conservative 93; Mismatches 207; Indels 79; Gaps 20;
 3 LSHAVSEDSVHSTFASRYVTSLPRFKQKPNISPKGAA-YQIINDEMLDGNPRNLIA 61
 77 LNEVKDEKSLRIVDRSPYEYFTTIPSHSVGRTVLRLLAATYDDEGPAFLG--RVSGA 134
 62 SF-----VTTWMEPECDKLMSSINKNYVDMDBYPTTELQRCVYMI 104
 135 VFNREDDKDEREMEEVFGKFAWTPLWPKL-----FPQVRIWEAEVVRMC 180
 105 AHLFNAPLEAEATAVGVTVGSSEAIMLAGLAFKRWKQKKAEGKPKVDKPNIVGTANVQ 164
 181 CNMNG--DSETC-GTWTGSGISILLACLAR-----NRLKRGKYTE--MIVPSSVH 230
 165 VCWEKFAFYFVELKEVKLSBEGYVMDPQAQVDMVDENTICVADILGSLN--GEFEDV 221
 231 AAFKAAECFRKVRKIPVDFVTFKVDLVKMAAINKRT-CM--LVGSAPNFPFGTVDDI 287

QY 222 KLLNDLLVEKNKETGWDTPHVDAAAGGFIAPFL-YPELEWDFRLPLVKSINVSCHKYGL 280
 DB 288 EAIGQLGLE-----YDIPVHVDAICLGGFLPFLEDEIRYDFRVPGVSSISADSHKYL 341
 QY 281 VYAGIGWVIRWKNEDLPBELIHPHINVLGADQPTFLNFSKSSQVIAQYYQLIRLGHGY 340
 DB 342 APKGSVLYRNKELLHNQYFCDAWQGGIYASATMEGSRAGHNIALCWAAMLXHAQEGY 401
 QY 341 ----RNVMNCRENNMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSCHTEFEISDMLR 396
 DB 402 KANARKIVDTRK----IRNGLSNIKIGIKLOGPSP--VCIVSWTTNDGV--ELYRPHFMK 454
 QY 397 RYGVWVPAVMPNQAQHTVLRVWIREDFSRTLAERLVIDIEKVMRELDLPSRVHKIS 456
 DB 455 EKHQWLNGLQFPAGV-HI-----MVTMNHHTHPLGLAEAFVADCRAAVEFVKS-----HKPS 503
 QY 457 LGQEKSES 464
 DB 504 ESDKTSEA 511

RESULT 9
 US-08-939-309-2
 ; Sequence 2, Application US/0899309
 ; Patent No. 6423527
 ; GENERAL INFORMATION:
 ; APPLICANT: Saba, Julie D.
 ; APPLICANT: Zhou, Jianhui
 ; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 ; TITLE OF INVENTION: POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 ; TITLE OF INVENTION: METHODS OF USE THEREFOR
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION NUMBER: US/08/939,309
 ; FILING DATE: 29-SEP-1997
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: David, Maki J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 200116.402
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 568 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-939-309-2
 Query Match 9.4%; Score 246.5; DB 4; Length 568;
 Best Local Similarity 23.2%; Pred. No. 5,2e-16;
 Matches 100; Conservative 76; Mismatches 210; Indels 45; Gaps 12;
 QY 22 YVTSLSLPRFKQKPNISPKGAA-YQIINDEMLDGNPRNLIAFPVTTWMEPECDKLMSSIN 81
 DB 112 YVKT-----LPAQMGTAELVRLERLKEYSGMDGQEGKASGAVNGPDKLTELIVQAYG 165
 QY 82 K---NVYDMEYPTTELQRCVYMIHLFNAPLEAEATAVGVTVGSSEAIMLAGLAF 137

```

166 EFTWSNPLHPDIIPGLRLEAEIVRMTCSLFLNG---GPDSCGCVTSGGTESILMACKAY 221
138 KRKQKRAEGKVPDKPNIVGANVQVCEKPARFVELKEVKLSEGYVMDPQQAVD 197
222 -----RDLALEKGIKTPETVAPESAHAFAFKAHYFGMKIVRVALKKNMEV-DYQAMKR 274
198 MVDENTICVADILGSLNGEFEDVKNLNDLLVEKNETGWDTPHVDAAASGGFIAPFL- 255
275 AISRNTAMLCVSTPQPHGVMDFVPEAKLTVR-----YKPLHVDACLGGFLVFMWEK 328
256 --Yp-ELEWDFRLPLVYKSNVSHGKYLIVYAGIWMVWRNKEDLPEELIFHINYLGADQP 312
329 AGYPLEKPFDRYKGVTSISADTHKYGYAPKSSVVMYSNEKYRTYQFFVGMADMGGYVA 388
313 TFLNFKGSSQVIAQYQIIRLHGHEGVRNWCENMIVLREGLEKTERENIVSKDBG 372
389 SPSIAGSRFGGIIAACWAAALMHFGENGIVEATKQIIKTARFLKSELENKKNIFIFG-DPQ 447
373 VPLVAFSLKDSCHTEFEISDMLRRYGVIVPAYTMPPPNAQHITVLRVIREDFSRFLAER 432
448 LSVIALGNSDFDI---YRLSNMWSAKGNWF-NYLPQPPRSIHFCITLHVHR-----KR 495
433 LVVIDIEKVMRE 443
496 VAIQFLKDIRE 506

```

```

RESULT 10
3-09-849-180-2
Sequence 2, Application US/09849180
Patent No. 6495359
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
                Zhou, Jianhui
TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98055
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/849,180
FILING DATE: 04-May-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Pepe, Jeffrey C.
REGISTRATION NUMBER: 46,985
REFERENCE/DOCKET NUMBER: 200116.402
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 692-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 568 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
-09-849-180-2
Query Match 9.4%; Score 246.5; DB 4; Length 568;
Best Local Similarity 23.2%; Pred. No. 5.2e-16;

```

```

Matches 100; Conservative 76; Mismatches 210; Indels 45; Gaps 12;
Qy 22 YVTSLSRFRKMPENSIPKEAAAYQIINDELMLDGNRFLNLASVVTWMPECDKLMSSIN 81
Db 112 YVKT-----LPAQGMGTAEVLERLREKYSMDGSGWQEGKASGAVYNGEPKLLELLVQAYG 165
Qy 82 K---NYVDMDEYVVTTELQNRQVNMIAHLFNAPLEEAETAVGVTVGSSSEAIMLAGLAF 137
Db 166 EFTWSNPLHPDIIPGLRLEAEIVRMTCSLFLNG---GPDSCGCVTSGGTESILMACKAY 221
Qy 138 KRKQKRAEGKVPDKPNIVGANVQVCEKPARFVELKEVKLSEGYVMDPQQAVD 197
Db 222 -----RDLALEKGIKTPETVAPESAHAFAFKAHYFGMKIVRVALKKNMEV-DYQAMKR 274
Qy 198 MVDENTICVADILGSLNGEFEDVKNLNDLLVEKNETGWDTPHVDAAASGGFIAPFL- 255
Db 275 AISRNTAMLCVSTPQPHGVMDFVPEAKLTVR-----YKPLHVDACLGGFLVFMWEK 328
Qy 256 --Yp-ELEWDFRLPLVYKSNVSHGKYLIVYAGIWMVWRNKEDLPEELIFHINYLGADQP 312
Db 329 AGYPLEKPFDRYKGVTSISADTHKYGYAPKSSVVMYSNEKYRTYQFFVGMADMGGYVA 388
Qy 313 TFLNFKGSSQVIAQYQIIRLHGHEGVRNWCENMIVLREGLEKTERENIVSKDBG 372
Db 389 SPSIAGSRFGGIIAACWAAALMHFGENGIVEATKQIIKTARFLKSELENKKNIFIFG-DPQ 447
Qy 373 VPLVAFSLKDSCHTEFEISDMLRRYGVIVPAYTMPPPNAQHITVLRVIREDFSRFLAER 432
Db 448 LSVIALGNSDFDI---YRLSNMWSAKGNWF-NYLPQPPRSIHFCITLHVHR-----KR 495
Qy 433 LVVIDIEKVMRE 443
Db 496 VAIQFLKDIRE 506

```

```

US-09-356-643B-6
; Sequence 6, Application US/09356643B
; Patent No. 6569666
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; POLYNUCLEOTIDES AND MODULATING AGENTS AND
; METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402CL
; CURRENT APPLICATION NUMBER: US/09/356,643B
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-356-643B-6
Query Match 9.4%; Score 246.5; DB 4; Length 568;
Best Local Similarity 23.2%; Pred. No. 5.2e-16;
Matches 100; Conservative 76; Mismatches 210; Indels 45; Gaps 12;
Qy 22 YVTSLSRFRKMPENSIPKEAAAYQIINDELMLDGNRFLNLASVVTWMPECDKLMSSIN 81
Db 112 YVKT-----LPAQGMGTAEVLERLREKYSMDGSGWQEGKASGAVYNGEPKLLELLVQAYG 165
Qy 82 K---NYVDMDEYVVTTELQNRQVNMIAHLFNAPLEEAETAVGVTVGSSSEAIMLAGLAF 137
Db 166 EFTWSNPLHPDIIPGLRLEAEIVRMTCSLFLNG---GPDSCGCVTSGGTESILMACKAY 221
Qy 138 KRKQKRAEGKVPDKPNIVGANVQVCEKPARFVELKEVKLSEGYVMDPQQAVD 197
Db 222 -----RDLALEKGIKTPETVAPESAHAFAFKAHYFGMKIVRVALKKNMEV-DYQAMKR 274
Qy 198 MVDENTICVADILGSLNGEFEDVKNLNDLLVEKNETGWDTPHVDAAASGGFIAPFL- 255

```

275 AISRNWMLVSTPDPFGVMDPVEVAKLIVR-----YKILPHVDAVLGGFLIVPMEX 328
 256 --YP-ELEWDFRLPIVAVKINVSCHKYGLVYAGIGWVIRNKEDLPEELIHFHINYLGDQP 312
 329 AGYPLKPKFDFRVKGVTSADTHKYGVAPKSSVWVMSNEKYTYQFFVGADWQGGVYA 388
 313 TFTLNFSGSSQVIAQYQLIRLHGHEGRVNMENCRNMIIVLREGLEKTERFNIVSKDEG 372
 389 SPSIAGSRFGIILAAACMAALMHFGVGEVATKQIITAREFKSELENIKNIIFIG-DPQ 447
 373 VPLVAFSLKDSSTCHTEFEISDMLRRYGIVPAYTPNNAQHITVLRVIREDFSRILAER 432
 448 LSVIALGSDNDFI---YRLSNMWSAKGNWF-NYLOFPSRHFCTILVHTR-----KR 495
 433 LVIDIEKVMRE 443
 496 VAIQFLKDOIRE 506

RESULT 12
 -08-939-309-4
 Sequence 4, Application US/08939309
 Patent No. 6423527
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 APPLICANT: Zhou, Jianhui
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 TITLE OF INVENTION: POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 TITLE OF INVENTION: METHODS OF USE THEREFOR
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/939,309
 FILING DATE: 29-SEP-1997
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: David, Maki J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 200116.402
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 -08-939-309-4

Query Match 8.7%; Score 227; DB 4; Length 568;
 Best Local Similarity 21.6%; Pred. No. 4.9e-14;
 Matches 98; Conservative 84; Mismatches 212; Indels 60; Gaps 14;
 8 SEBDSVSHSTFA---SRVYRTSLPRFKPENSIPKEAAAYQIINDELMLDGNPRLNLSFV 64
 95 TKDDISKWMSFLKVDKEYVKA-----LPSQGLSSAVLEKLEKEYSSMDAFWQGRASGT 148
 65 FTWMEPECDKLIIMSN-----KQVYMDDEYPTVTELRQNCVNMIAHLFNAPLEBAFVAG 120
 149 VYSGBEKTELLVAVGDFPAMSNPLHPIIFGLRKRKIEAIVRIACSLFNG----GPDSCG 204

121 VGVGSEAIMLA-----GLAFKRKQNKRAEGRKVPDKPNIVTGANVQVCWEKFAFYEE 175
 205 CVTSGGTESILMACKACRDLAFE-----KGIKTPVAVAPQSAHAAFNKAASYFG 253
 176 VELKEYKLGSEYVMDPQOAVMDVENTICVADILGSTLNGFEFVLDKLLNDLLVEKNRET 235
 254 MKIVRVPFLTKMEV-DVRAMRAISRNTAMLYCSTPQFPFHGVIDPVEYAKLAVK----- 307
 236 GMDTPHVDAAAGGFIAPEL-----YP-ELEWDFRPLVKSINVSCHKYGLVYAGIGWVIV 290
 308 -YKILPHVDAVLGGFLIVPMEXKAGYPLEHPFRVKGVTISADTHKYGVAPKSSGLVLY 366
 291 RNKEDLPEELIHFHINYLGDQFTFLNFKSGSSQVIAQYQLIRLHGHEGRVNMENCRN 350
 367 SDKKYENYQFFVTDWQGGIYASPTIAGSRPGGISAACWAAALMHFGVGEVATKQIIKT 426
 351 MIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSTCHTEFEISDMLRRYGIVPAYTPN 410
 427 ARFLKSELENIKGIFFVGNPQ-LSLIALGSRDFDI---YRLSNLMTAKGNLNLQLOFPFS 482
 411 AQH-ITVLRVIREDFSRTLAERLVIDIEKVMRE 443
 483 IHFCITLLHA-----RKRVAIQFLKDIRE 506

RESULT 13
 US-09-849-180-4
 Sequence 4, Application US/09849180
 Patent No. 6495359
 GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 APPLICANT: Zhou, Jianhui
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
 POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
 METHODS OF USE THEREFOR
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seed Intellectual Property Law Group
 STREET: 701 Fifth Avenue, Suite 6300
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98055

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/849,180
 FILING DATE: 04-May-2001
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Pepe, Jeffrey C.
 REGISTRATION NUMBER: 46,985
 REFERENCE/DOCKET NUMBER: 200116.402
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-849-180-4

Query Match 8.7%; Score 227; DB 4; Length 568;
 Best Local Similarity 21.6%; Pred. No. 4.9e-14;
 Matches 98; Conservative 84; Mismatches 212; Indels 60; Gaps 14;
 8 SEBDSVSHSTFA---SRVYRTSLPRFKPENSIPKEAAAYQIINDELMLDGNPRLNLSFV 64
 95 TKDDISKWMSFLKVDKEYVKA-----LPSQGLSSAVLEKLEKEYSSMDAFWQGRASGT 148
 65 FTWMEPECDKLIIMSN-----KQVYMDDEYPTVTELRQNCVNMIAHLFNAPLEBAFVAG 120
 149 VYSGBEKTELLVAVGDFPAMSNPLHPIIFGLRKRKIEAIVRIACSLFNG----GPDSCG 204

```

95 TKDDISKMWFLKVDKEYVKA-----LPSQGLSSSAVLEKLEKEYSSMDAFWQEGRASGT 148
65 TTWMEPCDKLIMSSIN-----KNYVDMDEYFVTELEQRCVNMIAHLFNAPLEAEATAVG 120
149 VYSGBEKLTELLKAYGDFAWSNPLHPDIFPGLRKTAEIVRIACSLFNG-----GPDSCG 204
121 VGTVGSSEAIMLA-----GLAFKRWKQKKAEGKPKVDKPNIVTGANVQVCWEKARYPE 175
205 CVTSGGTESILMACRDLAFE-----KGIKTPEIAPAQSAHAAFNKAASYFG 253
176 VELKEVKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEFEDYKLANDLLVEKNKET 235
254 MKIVRPLTKMVE-DVRAMRAISRNTAMLVGSTPQPHGVDPVPEVAKLAVK----- 307
236 GWDTPIHVDAASGGFIAPFL-----YP-ELEWDFRLPLVKSINVSCHKYGLVYAGIGWIW 290
308 -YKIFLHVDAKGLGFLIVFWEKAGYPLEHPDFRVRKGVTSISADTHKYGYAPKGSLSLVLY 366
291 RNKEDLPEELIFHINYLADQPTFTLNFSGSSQVIAQYQYQIRLGHGEGYRNVMNCREN 350
367 SDKXRYNQFFVDTMGGIYASPTIAGSRPGGISAACWAALMHFGENGYVEATKQIIKT 426
351 MIVLREGLEKTERNIVSKDEGVPVAFSLKDSCHTEFEISDMLRRYGWIYDXTMPPN 410
427 ARFLKSELENIKGIFVGNPQ-LSLIALGSRDFDI---YRLSNLMTAKGNLNLQLOFPF 482
411 AQH-ITVLRVIREDFSRTLAERLVIDIEKVMRE 443
483 IHFCITLLHA-----RKRVAIQFLKDIRE 506

```

```

RESULT 14
5-09-356-643B-8
Sequence 8, Application US/09356643B
Patent No. 6569666
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
TITLE OF INVENTION: METHODS OF USE THEREFOR
FILE REFERENCE: 200116.402C1
CURRENT APPLICATION NUMBER: US/09/356.643B
CURRENT FILING DATE: 1999-07-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 568
TYPE: PRT
ORGANISM: Homo sapiens
-09-356-643B-8
Query Match 8.7%; Score 227; DB 4; Length 568;
Best Local Similarity 21.6%; Pred. No. 4.9e-14;
Matches 98; Conservative 84; Mismatches 212; Indels 60; Gaps 14;
8 SESDVSVHSTFA---SRVYRTSLPRFKMPENSIPKEAAVQIINDELMLDGNRLNLASV 64
95 TKDDISKMWFLKVDKEYVKA-----LPSQGLSSSAVLEKLEKEYSSMDAFWQEGRASGT 148
65 TTWMEPCDKLIMSSIN-----KNYVDMDEYFVTELEQRCVNMIAHLFNAPLEAEATAVG 120
149 VYSGBEKLTELLKAYGDFAWSNPLHPDIFPGLRKTAEIVRIACSLFNG-----GPDSCG 204
121 VGTVGSSEAIMLA-----GLAFKRWKQKKAEGKPKVDKPNIVTGANVQVCWEKARYPE 175
205 CVTSGGTESILMACRDLAFE-----KGIKTPEIAPAQSAHAAFNKAASYFG 253
176 VELKEVKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEFEDYKLANDLLVEKNKET 235
254 MKIVRPLTKMVE-DVRAMRAISRNTAMLVGSTPQPHGVDPVPEVAKLAVK----- 307
236 GWDTPIHVDAASGGFIAPFL-----YP-ELEWDFRLPLVKSINVSCHKYGLVYAGIGWIW 290

```

```

308 -YKIFLHVDAKGLGFLIVFWEKAGYPLEHPDFRVRKGVTSISADTHKYGYAPKGSLSLVLY 366
291 RNKEDLPEELIFHINYLADQPTFTLNFSGSSQVIAQYQYQIRLGHGEGYRNVMNCREN 350
367 SDKXRYNQFFVDTMGGIYASPTIAGSRPGGISAACWAALMHFGENGYVEATKQIIKT 426
351 MIVLREGLEKTERNIVSKDEGVPVAFSLKDSCHTEFEISDMLRRYGWIYDXTMPPN 410
427 ARFLKSELENIKGIFVGNPQ-LSLIALGSRDFDI---YRLSNLMTAKGNLNLQLOFPF 482
411 AQH-ITVLRVIREDFSRTLAERLVIDIEKVMRE 443
483 IHFCITLLHA-----RKRVAIQFLKDIRE 506

```

```

RESULT 15
US-09-238-373-2
; Sequence 2, Application US/09238373A
; Patent No. 6187562
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID MALCOLM
; APPLICANT: GODDEN, ROBERT JAMES
; APPLICANT: TESTA, TANIA TAMSON
; TITLE OF INVENTION: NOVEL COMPOUNDS
; CURRENT APPLICATION NUMBER: US/09/238,373A
; CURRENT FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: UK 9824026.0
; EARLIER FILING DATE: 1998-11-03
; EARLIER APPLICATION NUMBER: EP 98300625.5
; EARLIER FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-238-373-2
Query Match 8.6%; Score 224; DB 3; Length 568;
Best Local Similarity 21.4%; Pred. No. 9.9e-14;
Matches 96; Conservative 86; Mismatches 217; Indels 50; Gaps 13;
QY 8 SESDVSVHSTFA---SRVYRTSLPRFKMPENSIPKEAAVQIINDELMLDGNRLNLASV 64
DB 95 TKDDISKMWFLKVDKEYVKA-----LPSQGLSSSAVLEKLEKEYSSMDAFWQEGRASGT 148
QY 65 TTWMEPCDKLIMSSIN-----KNYVDMDEYFVTELEQRCVNMIAHLFNAPLEAEATAVG 120
DB 149 VYSGBEKLTELLKAYGDFAWSNPLHPDIFPGLRKTAEIVRIACSLFNG-----GPDSCG 204
QY 121 VGTVGSSEAIMLAGLAFKRWKQKKAEGKPKVDKPNIVTGANVQVCWEKARYPEVELKE 180
DB 205 CVTSGGTESILMACRDLAFE-----KGIKTPEIAPAQSAHAAFNKAASYFGMKIVR 258
QY 181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEFEDYKLANDLLVEKNKETGWDTP 240
DB 259 VPLTKMVE-DVRAMRAISRNTAMLVGSTPQPHGVDPVPEVAKLAVK-----YKIP 311
QY 241 IHVDAASGGFIAPFL-----YP-ELEWDFRLPLVKSINVSCHKYGLVYAGIGWIWRNKED 295
DB 312 LHVDAKGLGFLIVFWEKAGYPLEHPDFRVRKGVTSISADTHKYGYAPKGSLSLVLYSDKY 371
QY 296 LPDELPHINYLADQPTFTLNFSGSSQVIAQYQYQIRLGHGEGYRNVMNCRENIVLR 355
DB 372 RNYQFFVDTMGGIYASPTIAGSRPGGISAACWAALMHFGENGYVEATKQIIKTARELK 431
QY 356 EGLEKTERNIVSKDEGVPVAFSLKDSCHTEFEISDMLRRYGWIYDXTMPPNAQH-I 414
DB 432 SELENIKGIFVGNPQ-LSVIALGSRDFDI---YRLSNLMTAKGNLNLQLOFPFPIHFCI 487
QY 415 TVLRVIREDFSRTLAERLVIDIEKVMRE 443

```


1 protein - protein search, using sw model
 GenCore version 5.1.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.
 October 22, 2003, 13:34:43 ; Search time 69 Seconds
 (without alignments)
 1154.793 Million cell updates/sec

US-10-006-852-2
 tle: rfact score: 2615
 quence: 1 MVLHSHVSESDVSHSTFAS.....DIITGKKFVADRKKTSIGIC 502
 oring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 arched: 1107863 seqs, 158726573 residues
 tal number of hits satisfying chosen parameters: 1107863
 nimum DB seq length: 0
 aximum DB seq length: 2000000000
 ost-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

atabase : A_Geneseq_19Jun03.*

1:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	2615	100.0	502	AAO15132	A thaliana GAD1.
2	2607	99.7	502	ABB93521	Herbicidally activ
3	2250	86.0	500	AAO15139	Petunia GAD1 protei
4	2240	85.7	496	AAO15137	Tobacco GAD1. Nic
5	2199	84.1	496	AAO15138	Tobacco GAD2. Nic
6	2135.5	81.7	493	AAO15135	A thaliana GAD4.
7	2135.5	81.7	493	ABB91695	Herbicidally activ
8	2130	81.5	494	AAO15133	A thaliana GAD2.
9	2130	81.5	494	ABB91465	Herbicidally activ

10	2080	79.5	500	23	AAO15134	A thaliana GAD3.
11	2080	79.5	500	23	ABB91694	Herbicidally activ
12	2014	77.0	502	23	AAO15140	Tomato GAD protein
13	1969	75.3	494	23	AAO15136	A thaliana GAD5.
14	1969	75.3	494	23	ABB92352	Herbicidally activ
15	1965	75.1	494	21	AAO15137	Arabidopsis thalia
16	1965	75.1	507	21	AAO15136	Arabidopsis thalia
17	1876.5	71.8	464	21	AAO15136	Arabidopsis thalia
18	1318	50.4	514	22	ABB96992	A oryzae glutamate
19	1188.5	45.4	479	21	AAO15136	Arabidopsis thalia
20	1175	44.9	464	21	AAO15136	Arabidopsis thalia
21	1050.5	40.2	430	21	AAO15136	Arabidopsis thalia
22	1028	39.3	466	23	ABB94627	Lactococcus lactis
23	1023	39.1	466	19	AAW59165	L. lactis MG3316 g
24	955.5	36.5	462	23	ABB47826	Listeria monocytog
25	953.5	36.5	462	23	ABB48689	Listeria monocytog
26	940	35.9	464	23	ABB49717	Listeria monocytog
27	697	26.7	1042	22	ABG17675	Novel human diagno
28	697	26.7	1042	22	ABG25740	Novel human diagno
29	626.5	24.0	194	23	ABB92351	Herbicidally activ
30	508.5	19.4	609	22	ABG17670	Novel human diagno
31	504	19.3	1155	22	ABG25739	Novel human diagno
32	504	19.3	1231	22	ABG17676	Novel human diagno
33	504	19.3	1231	22	ABG24397	Novel human diagno
34	473.5	18.1	572	22	ABG25741	Novel human diagno
35	393.5	15.0	191	22	AAU87540	Novel central nerv
36	393.5	15.0	191	22	AAU22935	Novel human enzyme
37	393.5	15.0	191	22	AAU18706	Renal and cardiova
38	343	13.1	213	22	ABG08876	Novel human diagno
39	325	12.4	589	20	AAV05629	Yeast sphingosine-
40	323	12.4	589	22	AAO15132	S cerevisiae apopt
41	314	12.0	384	22	AAO15132	Putative P. abyssi
42	304.5	11.6	589	22	AAO15132	C albicans apoptos
43	299.5	11.5	545	22	AAO15132	Drosophila melanog
44	299.5	11.5	545	22	AAO15132	Drosophila melanog
45	299.5	11.5	545	22	AAO15132	Drosophila melanog

ALIGNMENTS

RESULT 1
 AAO15132
 ID AAO15132 standard; Protein; 502 AA.
 AC AAO15132;
 DT 25-SEP-2002 (first entry)
 DE A thaliana GAD1.
 KW GAD; plant GABA production regulation; glutamic acid decarboxylase;
 KW plant stress; GABA; gamma-aminobutyric acid; stress resistance.
 OS Arabidopsis thaliana.
 FN WO200238736-A2.
 FD 16-MAY-2002.
 PF 07-NOV-2001; 2001WO-US47447.
 PR 07-NOV-2000; 2000US-246367P.
 XX (EMER-) EMERALD BIOAGRICULTURE CORP.
 XX Kinnersley AM, Turano FU;
 XX WPI; 2002-490073/52.
 XX N-PSDE; AAL43410.
 XX Making transformed plants that selectively increase gamma-aminobutyric
 XX acid production, by incorporating a DNA construct with a polynucleotide

I I encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 X X Claim 17; Page 53; 63pp; English.
 S S The present invention relates to a method of producing a transformed
 X X plant that selectively increases production of gamma-aminobutyric acid
 C C (GABA) in response to a signal, by incorporating into the plant's genome
 C C a DNA construct with a non-constitutive promoter operably linked to a
 C C polynucleotide encoding a functional plant glutamic acid decarboxylase
 C C (GAD), to provide a transformed plant that expresses the GAD coding
 C C sequence in response to a signal. Plants of this type have an enhanced
 C C ability to tolerate environmental or other stresses. The present sequence
 C C is the A. thaliana GAD1 protein.
 X X Sequence 502 AA;

Query Match 100.0%; Score 2615; DB 23; Length 502;
 Best Local Similarity 100.0%; Pred. No. 1.4e-241;
 Matches 502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 MVLSHAVSESDVSHSTFASRYVTSLPFRKMPENSIPKEAAYQIINDELMLDGNPRNL 60
 D 1 MVLSHAVSESDVSHSTFASRYVTSLPFRKMPENSIPKEAAYQIINDELMLDGNPRNL 60
 Y 61 ASFTVTWMEPECDKLISSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEEAFTAVG 120
 D 61 ASFTVTWMEPECDKLISSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEEAFTAVG 120
 Y 121 VGTVGSSEAIMLAGLAFKRWKQKRAEKPVDKPNIVTGANVQVCWEKFAFYFEVLEKE 180
 D 121 VGTVGSSEAIMLAGLAFKRWKQKRAEKPVDKPNIVTGANVQVCWEKFAFYFEVLEKE 180
 Y 181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLVEKNKGTGMDTP 240
 D 181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLVEKNKGTGMDTP 240
 Y 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRNKEDLPEEL 300
 D 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRNKEDLPEEL 300
 Y 301 IFHINYLGADQPTFLNFKSGSSQVIAQYYQLIRLGHGGRVNMENCRENMIVLREGLEK 360
 D 301 IFHINYLGADQPTFLNFKSGSSQVIAQYYQLIRLGHGGRVNMENCRENMIVLREGLEK 360
 Y 361 TERFNIVSKDEGVPLVAFSLKDSCHTEPEISDMLRRYGIWVPAVTMPNQAQHIIVLRVV 420
 D 361 TERFNIVSKDEGVPLVAFSLKDSCHTEPEISDMLRRYGIWVPAVTMPNQAQHIIVLRVV 420
 Y 421 IREDFSRTLAERLVLDIEKVMRELDLPSRVIHKISLGQEKSESNDMLMVTYKKSDDIDK 480
 D 421 IREDFSRTLAERLVLDIEKVMRELDLPSRVIHKISLGQEKSESNDMLMVTYKKSDDIDK 480
 Y 481 QRDIIITGWKKFVADRKKTSGC 502
 D 481 QRDIIITGWKKFVADRKKTSGC 502

RESULT 2
) ABB93521
) ABB93521 standard; Protein; 502 AA.
) ABB93521;
) 31-MAY-2002 (first entry)
) Herbicidally active polypeptide SEQ ID NO 2732.
) Herbicidal; plant; agriculture; herbicide.
) Arabidopsis thaliana.
) WO200210210-A2.

XX 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX (FARB) BAYER AG.
 XX Tietjen K, Weidler M;
 XX WPI; 2002-269010/31.
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX Claim 5; SEQ ID NO 2732; 261pp + Sequence Listing; English.
 XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX Sequence 502 AA;

Query Match 99.7%; Score 2607; DB 23; Length 502;
 Best Local Similarity 99.8%; Pred. No. 7.9e-241;
 Matches 501; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVLSHAVSESDVSHSTFASRYVTSLPFRKMPENSIPKEAAYQIINDELMLDGNPRNL 60
 Db 1 MVLSHAVSESDVSHSTFASRYVTSLPFRKMPENSIPKEAAYQIINDELMLDGNPRNL 60
 Qy 61 ASFTVTWMEPECDKLISSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEEAFTAVG 120
 Db 61 ASFTVTWMEPECDKLISSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEEAFTAVG 120
 Qy 121 VGTVGSSEAIMLAGLAFKRWKQKRAEKPVDKPNIVTGANVQVCWEKFAFYFEVLEKE 180
 Db 121 VGTVGSSEAIMLAGLAFKRWKQKRAEKPVDKPNIVTGANVQVCWEKFAFYFEVLEKE 180
 Qy 181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLVEKNKGTGMDTP 240
 Db 181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLVEKNKGTGMDTP 240
 Qy 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRNKEDLPEEL 300
 Db 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSCHKYGLVYAGIGWVIRNKEDLPEEL 300
 Qy 301 IFHINYLGADQPTFLNFKSGSSQVIAQYYQLIRLGHGGRVNMENCRENMIVLREGLEK 360
 Db 301 IFHINYLGADQPTFLNFKSGSSQVIAQYYQLIRLGHGGRVNMENCRENMIVLREGLEK 360
 Qy 361 TERFNIVSKDEGVPLVAFSLKDSCHTEPEISDMLRRYGIWVPAVTMPNQAQHIIVLRVV 420
 Db 361 TERFNIVSKDEGVPLVAFSLKDSCHTEPEISDMLRRYGIWVPAVTMPNQAQHIIVLRVV 420
 Qy 421 IREDFSRTLAERLVLDIEKVMRELDLPSRVIHKISLGQEKSESNDMLMVTYKKSDDIDK 480
 Db 421 IREDFSRTLAERLVLDIEKVMRELDLPSRVIHKISLGQEKSESNDMLMVTYKKSDDIDK 480
 Qy 481 QRDIIITGWKKFVADRKKTSGC 502
 Db 481 QRDIIITGWKKFVADRKKTSGC 502

RESULT 3
 X D AAO15139 standard; Protein; 500 AA.
 X C AAO15139;
 X Y 25-SEP-2002 (first entry)
 X X Petunia GAD protein.
 X X GAD; plant GABA production regulation; glutamic acid decarboxylase;
 X X plant stress; GABA; gamma-aminobutyric acid; stress resistance.
 X X Petunia sp.
 X X WO200238736-A2.
 X X 16-MAY-2002.
 X X 07-NOV-2001; 2001WO-US47447.
 X X 07-NOV-2000; 2000US-246367P.
 X X (EMER-) EMERALD BIOAGRICULTURE CORP.
 X X Kinnersley AM, Turano FJ;
 X X MPI; 2002-490073/52.
 X X N-PSDB; AAL43417.
 X X Making transformed plants that selectively increase gamma-aminobutyric
 X X acid production, by incorporating a DNA construct with a polynucleotide
 X X encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 X X
 X X Claim 17; Page 58; 63pp; English.
 X X The present invention relates to a method of producing a transformed
 X X plant that selectively increases production of gamma-aminobutyric acid
 X X (GABA) in response to a signal, by incorporating into the plant's genome
 X X a DNA construct with a non-constitutive promoter operably linked to a
 X X polynucleotide encoding a functional plant glutamic acid decarboxylase
 X X (GAD), to provide a transformed plant that expresses the GAD coding
 X X sequence in response to a signal. Plants of this type have an enhanced
 X X ability to tolerate environmental or other stresses. The present sequence
 X X is the petunia GAD protein.
 X X Sequence 500 AA;
 X X Query Match 86.0%; Score 2250; DB 23; Length 500;
 X X Best Local Similarity 84.9%; Pred. No. 1.4e-206;
 X X Matches 427; Conservative 39; Mismatches 33; Indels 4; Gaps 2;
 X X
 X X 1 MVLSHVSESDVSVHSTFASRYVRTSLPRFKMPNSIPKEAAYQIINDELMLDGNPRLNL 60
 X X 1 MVLKTSVQSDVSIHSTFASRYVRTSLPRFKMPNSIPKEAAYQIINDELMLDGNPRLNL 60
 X X 61 ASFTVWNEPCDKLIMSSINKNVYDMDFPVTTELQRCVNNMIAHLFNAPLBEAETAVG 120
 X X 61 ASFTVWNEPCDKLIMDSINKNVYDMDFPVTTELQRCVNNMIAHLFNAPLBEGETAVG 120
 X X 121 VGTVGSSSAIMLAGLAFKRWQNKARQKGFVDPKNIIVTGANVQVCWEKFARYFEVELKE 180
 X X 121 VGTVGSSSAIMLAGLAFKRWQNKARQKGFVDPKNIIVTGANVQVCWEKFARYFEVELKE 180
 X X 181 VKLSEGYVYVMDPQAVDMVDENTICVADILGSTLNGEFEDVKLINDLLVEKNKETGMDTP 240
 X X 181 VKLSEGYVYVMDPERAVENVDENTICVAAILGSTLNGEFEDVKRINDLLVEKNKETGMDTP 240
 X X 241 IHVDAASGGFIAPFLYPLEWDFRPLIVKYSINVSGHKYGLVYAGIWMVWRNKEDLPEEL 300
 X X 241 IHVDAASGGFIAPFLYPLEWDFRPLIVKYSINVSGHKYGLVYAGIWMVWRNKODLPEEL 300

301 IFHINYLGADQPTFTLNFSKSSQVIAQYQYQIRLGHGEGYRNVMENKRENMIVLREGLEK 360
 301 IFHINYLGADQPTFTLNFSKSSQVIAQYQYQIRLGHGEGYRNVMENKRENMIVLREGLEK 360
 361 TERFNIVSKDEGVPLVAFSLKSDSSCHTEFEISDMRRYGWIVPAYTMPNQAHIITVLRVV 420
 361 TGRFNIIISKEIGVPLVAFSLKDNRQHFNEFISETLRRFGWIVPAYTMPNQAHIITVLRVV 420
 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVTHKISLGOEKSESNLDLMTVTKKSDIDK 480
 421 IREDFSRITLAERLVIDIEKVLHELDLTPARVNAKLAVAEQAANGSEVH---KFTDSEV 477
 481 QRDITIGWKKFVAD-RKKTSGIC 502
 478 OLEMITAWKKFVEEKKKTNRVC 500
 RESULT 4
 ID AAO15137 standard; Protein; 496 AA.
 AC AAO15137;
 DT 25-SEP-2002 (first entry)
 XX Tobacco GAD1.
 KW GAD; plant GABA production regulation; glutamic acid decarboxylase;
 KW plant stress; GABA; gamma-aminobutyric acid; stress resistance.
 XX Nicotiana tabacum.
 XX WO200238736-A2.
 XX 16-MAY-2002.
 XX 07-NOV-2001; 2001WO-US47447.
 XX 07-NOV-2000; 2000US-246367P.
 (EMER-) EMERALD BIOAGRICULTURE CORP.
 Kinnersley AM, Turano FJ;
 MPI; 2002-490073/52.
 N-PSDB; AAL43415.
 Making transformed plants that selectively increase gamma-aminobutyric
 acid production, by incorporating a DNA construct with a polynucleotide
 encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 Claim 17; Page 57; 63pp; English.
 The present invention relates to a method of producing a transformed
 plant that selectively increases production of gamma-aminobutyric acid
 (GABA) in response to a signal, by incorporating into the plant's genome
 a DNA construct with a non-constitutive promoter operably linked to a
 polynucleotide encoding a functional plant glutamic acid decarboxylase
 (GAD), to provide a transformed plant that expresses the GAD coding
 sequence in response to a signal. Plants of this type have an enhanced
 ability to tolerate environmental or other stresses. The present sequence
 is the tobacco GAD1 protein.
 Sequence 496 AA;
 Query Match 85.7%; Score 2240; DB 23; Length 496;
 Best Local Similarity 85.1%; Pred. No. 1.2e-205;
 Matches 428; Conservative 35; Mismatches 32; Indels 8; Gaps 3;
 1 MVLHAVSESDVSVHSTFASRYVRTSLPRFKMPNSIPKEAAYQIINDELMLDGNPRLNL 60
 1 MVLKTSVQSDVSIHSTFASRYVRTSLPRFKMPNSIPKEAAYQIINDELMLDGNPRLNL 60

(EMER-) EMERALD BIOAGRICULTURE CORP.

Kinnersley AM, Turano FJ;

WPI; 2002-490073/52.
N-ESDB; AAL43413.

Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome

Claim 17; Page 55-56; 63pp; English.

The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the A. thaliana GAD4 protein.

Sequence 493 AA;

Query Match 81.7%; Score 2135.5; DB 23; Length 493;
Best Local Similarity 82.4%; Pred. No. 1.3e-195;
Matches 411, Conservative 32; Mismatches 45; Indels 11; Gaps 2;

1 MVLSHAVSESDVSHSTFASRYVTSIPRFKMPENSIPKEAAAYQIINDELMLDGNPRLNL 60

1 MVLSTVSESDVSHSTFASRYVRSNLSRPFEMPENSIPKEAAAYQIINDELMLDGNPRLNL 60

61 ASFVTTWMEPCDKLIMSSINKNYVDMDEYVPTTELQNRVNMIAHLFNAPLEEAETAVG 120

61 ASFVTTWMEPCDKLIMSSINKNYVDMDEYVPTTELQNRVNMIAHLFNAPLEEAETAVG 120

121 VGTGSSSEAIMLAGLAFKRWQNKRAEGKVPDKPNIIVTGANVQVCWEKFARYFEVELKE 180

121 VGTGSSSEAIMLAGLAFKRWQNKRAEGKVPDKPNIIVTGANVQVCWEKFARYFEVELKE 180

181 VKLSEGVYVMDPQAAVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKKNKGTGWDTP 240

181 VNLREDIYVMDPVKAVEMVDENTICVAALIGSTLTGFEFEDVKLLNDLLVEKKNKGTGWDTP 240

241 IHVDAASGGFTAPFLPELEWDFRFLPLVKSINVSCHKYGLVYAGIGWVWVWRTKTDLPDEL 300

241 IHVDAASGGFTAPFLPELEWDFRFLPLVKSINVSCHKYGLVYAGIGWVWVWRTKTDLPDEL 300

301 IFHINYLGAQOPTFTLNFSGSSQVIAQYQLIRLGHGEGYRWNTMNCRENMIIVLREGLEK 360

301 IFHINYLGAQOPTFTLNFSGSSQVIAQYQLIRLGHGEGYRWNTMNCRENMIIVLREGLEK 360

361 TERFNIVSKDGVPLVAFSLKDSCHTEFEISDMLRRYGVIVPAYTMPNAQHITVLRVV 420

361 TGRFKIVSKENGVPVAFSLKDSCHTEFEISDMLRRYGVIVPAYTMPNAQHITVLRVV 420

421 IREDFSRVLAERLVADEKVLHELDLTPARVHAKVANGKVG-----VKKTPET 470

421 IREDFSRVLAERLVADEKVLHELDLTPARVHAKVANGKVG-----VKKTPET 470

481 ORDITGKWKVADRKKTS 499

471 QREVTAWKLL-ETKKTIN 488

RESULT 7

ABB91695

ID ABB91695 standard; Protein; 493 AA.

XX

31-MAY-2002 (first entry)

Herbicide active polypeptide SEQ ID NO 906.

Herbicide; plant; agriculture; herbicide.

Arabicopsis thaliana.

WO2000210210-A2.

XX 31-MAY-2002 (first entry)
XX Herbicide active polypeptide SEQ ID NO 906.
XX Herbicide; plant; agriculture; herbicide.
XX Arabicopsis thaliana.
XX WO2000210210-A2.
XX 07-FEB-2002.
XX 28-AUG-2001; 2001WO-EP09892.
XX 28-AUG-2001; 2001WO-EP09892.
XX (FARB) BAYER AG.
XX Tietjen K, Weidler M;
XX WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms -
XX Claim 5; SEQ ID NO 906; 261pp + Sequence Listing; English.
XX The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
XX with nucleic acid or amino acid sequences from non-plant organisms using
XX suitable search parameters, where plant sequences having an E-value
XX greater by a factor of 3 than the E-value of most similar non-plant
XX sequences are selected. The polypeptides or nucleic acids encoding them
XX are useful for identifying modulators. The identified modulators are
XX useful as herbicides.
XX Sequence 493 AA;
SQ Query Match 81.7%; Score 2135.5; DB 23; Length 493;
Best Local Similarity 82.4%; Pred. No. 1.3e-195;
Matches 411, Conservative 32; Mismatches 45; Indels 11; Gaps 2;
QY 1 MVLSHAVSESDVSHSTFASRYVTSIPRFKMPENSIPKEAAAYQIINDELMLDGNPRLNL 60
Db 1 MVLSTVSESDVSHSTFASRYVRSNLSRPFEMPENSIPKEAAAYQIINDELMLDGNPRLNL 60
QY 61 ASFVTTWMEPCDKLIMSSINKNYVDMDEYVPTTELQNRVNMIAHLFNAPLEEAETAVG 120
Db 61 ASFVTTWMEPCDKLIMSSINKNYVDMDEYVPTTELQNRVNMIAHLFNAPLEEAETAVG 120
QY 121 VGTGSSSEAIMLAGLAFKRWQNKRAEGKVPDKPNIIVTGANVQVCWEKFARYFEVELKE 180
Db 121 VGTGSSSEAIMLAGLAFKRWQNKRAEGKVPDKPNIIVTGANVQVCWEKFARYFEVELKE 180
QY 181 VKLSEGVYVMDPQAAVDMVDENTICVADILGSLTNGEFEDVKLLNDLLVEKKNKGTGWDTP 240
Db 181 VNLREDIYVMDPVKAVEMVDENTICVAALIGSTLTGFEFEDVKLLNDLLVEKKNKGTGWDTP 240
QY 241 IHVDAASGGFTAPFLPELEWDFRFLPLVKSINVSCHKYGLVYAGIGWVWVWRTKTDLPDEL 300
Db 241 IHVDAASGGFTAPFLPELEWDFRFLPLVKSINVSCHKYGLVYAGIGWVWVWRTKTDLPDEL 300
QY 301 IFHINYLGAQOPTFTLNFSGSSQVIAQYQLIRLGHGEGYRWNTMNCRENMIIVLREGLEK 360
Db 301 IFHINYLGAQOPTFTLNFSGSSQVIAQYQLIRLGHGEGYRWNTMNCRENMIIVLREGLEK 360
QY 361 TERFNIVSKDGVPLVAFSLKDSCHTEFEISDMLRRYGVIVPAYTMPNAQHITVLRVV 420
Db 361 TGRFKIVSKENGVPVAFSLKDSCHTEFEISDMLRRYGVIVPAYTMPNAQHITVLRVV 420

421 IREDFSRRLAERLVLDIEKVMRELDLPSRVHKLISLGOEKSSNSDNLMMVTVKKSIDDK 480
 421 IREDFSRRLAERLVADFEKVLHELDLPLRVHAKVANGKNG-----VKKTFEET 470
 481 QRDIIITGWKKFVADRKKTS 499
 471 QREVTAYWKKLL-ETKKTN 488

SULT 8
 O15133
 AAO15133 standard; Protein; 494 AA.
 AAO15133;
 25-SEP-2002 (first entry)
 A thaliana GAD2.
 GAD; plant GABA production regulation; glutamic acid decarboxylase;
 plant stress; GABA; gamma-aminobutyric acid; stress resistance.
 Arabidopsis thaliana.
 WO200238736-A2.
 16-MAY-2002.
 07-NOV-2001; 2001WO-US47447.
 07-NOV-2000; 2000US-246367P.
 (EMER-) EMERALD BIOAGRICULTURE CORP.
 Kinnersley AM, Turano FJ;
 WPI: 2002-490073/52.
 N-PSDB; AA143411.
 Making transformed plants that selectively increase gamma-aminobutyric
 acid production, by incorporating a DNA construct with a polynucleotide
 encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 -
 Claim 17; Page 54; 63pp; English.
 The present invention relates to a method of producing a transformed
 plant that selectively increases production of gamma-aminobutyric acid
 (GABA) in response to a signal, by incorporating into the plant's genome
 a DNA construct with a non-constitutive promoter operably linked to a
 polynucleotide encoding a functional plant glutamic acid decarboxylase
 (GAD), to provide a transformed plant that expresses the GAD coding
 sequence in response to a signal. Plants of this type have an enhanced
 ability to tolerate environmental or other stresses. The present sequence
 is the A. thaliana GAD2 protein.
 Sequence 494 AA;
 Query Match 81.5%; Score 2130; DB 23; Length 494;
 Best Local Similarity 80.2%; Pred. No. 4.4e-195; Mismatches 42; Indels 12; Gaps 3;
 Matches 404; Conservativity 46; Mismatches 42; Indels 12; Gaps 3;

Db 120 VGTGSSSEIAMLGLAFKRWKQKRKABGKPYDKPNIVTGANVQVCWEKFPARYFEVELKE 179
 QY 181 VKLSEGYVMDPQAVDMVDENTICVADILGSTLNGEPEDVKLLDNDLLVBNKNETGWPTP 240
 Db 180 VNLSEGYVMDPDKAAEMVDENTICVAAILGSTLNGEPEDVKLLDNDLLVBNKNETGWPTP 239
 QY 241 IHVDAASGGFIAPLPELWDFRLPLVKSINVSQHKYGLVYAGIGWIVWRNKNKEDLPEEL 300
 Db 240 IHVDAASGGFIAPLPELWDFRLPLVKSINVSQHKYGLVYAGIGWIVWRNKNKEDLPEEL 299
 QY 301 IFHINYLGAQDPTFTLNFSGSSQVIAQYQLIRLGHGEGYRNVMENCRNMTVLRGLGK 360
 Db 300 IFHINYLGAQDPTFTLNFSGSSQVIAQYQLIRLGHGEGYRNVMENCRNMTVLRGLGK 359
 QY 361 TERNIVSKDQGVVAFSLKDSCHTEPEISDMLRRYGVIPAYTMPNACHITVLRVV 420
 Db 360 TERNIVSKDQGVVAFSLKDSCHTEPEISDMLRRYGVIPAYTMPNACHITVLRVV 419
 QY 421 IREDFSRRLAERLVLDIEKVMRELDLPSRVHKLISLGOEKSSNSDNLMMVTVKKSIDDK 480
 Db 420 IREDFSRRLAERLVADFEKVLHELDLPLRVHAKVANGKNG-----VKKTFEET 470
 QY 481 Q--RDIITGWKKFVADRKKTS 502
 Db 471 EILMEVIVGWRKFKVKKRKGNGVC 494

RESULT 9
 ABB91465
 ID ABB91465 standard; Protein; 494 AA.
 XX ABB91465;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 676.
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EF09892.
 XX
 PR 28-AUG-2001; 2001WO-EF09892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 organisms -
 PT
 PS Claim 5; SEQ ID NO 676; 261pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX
 SQ Sequence 494 AA;

Query Match 81.5%; Score 2130; DB 23; Length 494;
 Best Local Similarity 80.2%; Pred. No. 4.4e-195;
 Matches 404; Conservative 46; Mismatches 42; Indels 12; Gaps 3;

1 MVLSHAVSSEVSVHSTFASRYVRTSLPFRKMPENSIPKAAAYQIINDELMLDGNPRNL 60
 1 MWLTKTAT-NDSVCTMFGSRVVRTSLPKYIENGNSIPKDAAYQIIKDELMLDGNPRNL 59
 61 ASFVTTWMEPCDCKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEFTAVG 120
 60 ASFVTTWMEPCDCKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEFTAVG 119
 121 VGTVGSSEAIMLAGLAFKRWKQKKAQKPKVDKFNIVTGANVQVCWEKFAFYFEVLEKE 180
 120 VGTVGSSEAIMLAGLAFKRWKQKKAQKPKVDKFNIVTGANVQVCWEKFAFYFEVLEKE 179
 181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKKTGWDT 240
 180 VNLSEGYVMDPQQAQVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKKTGWNT 239
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 240 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 299
 301 IFHINYLGADQPTFLNFSKSSQVIAQYQLIRLGHGEGYRNVMENCRENMIVLREGLEK 360
 300 IFHINYLGADQPTFLNFSKSSQVIAQYQLIRLGHGEGYRNVMENCIENMVVLEKIEK 359
 361 TERFNIVSKDEGVPLVAFSLKDSCHTEFEISDLRRYGWIVPAYTPPNAQHIITVLRVV 420
 360 TERFNIVSKDQGVPLVAFSLKDSCHTEFEISDLRRYGWIVPAYTPPNAQHIITVLRVV 419
 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGQESKSENSDMLVTVKSDIDK 480
 420 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGQESKSENSDMLVTVKSDIDK 479
 481 Q--RDIIITGKVKVADRKTSTGI 502
 471 EILMEVIVGWKRFVRRKKKQNGVC 494

RESULT 10
 AAO15134
 ID AAO15134 standard; Protein; 500 AA.
 AC AAO15134;
 XX XX
 DT 25-SEP-2002 (first entry)
 XX XX
 DE A thaliana GAD3.
 XX XX
 KW GAD; plant GABA production regulation; glutamic acid decarboxylase;
 KW plant stress; GABA; gamma-aminobutyric acid; stress resistance.
 OS Arabidopsis thaliana.
 XX XX
 PN WO200238736-A2.
 PD 16-MAY-2002.
 PF 07-NOV-2001; 2001WO-US47447.
 XX XX
 PR 07-NOV-2000; 2000US-246367P.
 XX XX
 PA (EMER-) EMERALD BIOAGRICULTURE CORP.
 XX XX
 PI Kinnersley AM, Turano FU;
 XX XX
 DR WPT; 2002-490073/52.
 DR N-PSDB; AAL43412.
 XX XX
 PT Making transformed plants that selectively increase gamma-aminobutyric

PT acid production, by incorporating a DNA construct with a polynucleotide
 PT encoding a plant glutamic acid decarboxylase enzyme into plant's genome
 XX
 XX
 PS Claim 17; Page 55; 63pp; English.
 CC The present invention relates to a method of producing a transformed
 CC plant that selectively increases production of gamma-aminobutyric acid
 CC (GABA) in response to a signal, by incorporating into the plant's genome
 CC a DNA construct with a non-constitutive promoter operably linked to a
 CC polynucleotide encoding a functional plant glutamic acid decarboxylase
 CC (GAD), to provide a transformed plant that expresses the GAD coding
 CC sequence in response to a signal. Plants of this type have an enhanced
 CC ability to tolerate environmental or other stresses. The present sequence
 CC is the A. thaliana GAD1 protein.
 SQ Sequence 500 AA;
 Query Match 79.5%; Score 2080; DB 23; Length 500;
 Best Local Similarity 78.8%; Pred. No. 2.8e-190;
 Matches 395; Conservative 44; Mismatches 52; Indels 10; Gaps 1;

1 MVLGHAVSESDVSVHSTFASRYVRTSLPFRKMPENSIPKAAAYQIINDELMLDGNPRNL 60
 1 MVLSEKASKSDSSIHSTFASRYVRTSLPFRKMPENSIPKAAAYQIINDELMLDGNPRNL 60
 61 ASFVTTWMEPCDCKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEFTAVG 120
 61 ASFVTTWMEPCDCKLIMSSINKNYVDMDEYPTTELQNRQVNMIAHLFNAPLEAEFTAVG 120
 121 VGTVGSSEAIMLAGLAFKRWKQKKAQKPKVDKFNIVTGANVQVCWEKFAFYFEVLEKE 180
 121 VGTVGSSEAIMLAGLAFKRWKQKKAQKPKVDKFNIVTGANVQVCWEKFAFYFEVLEKE 180
 181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKKTGWDT 240
 181 VKLSEGYVMDPQQAQVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKKTGWDT 240
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 301 IFHINYLGADQPTFLNFSKSSQVIAQYQLIRLGHGEGYRNVMENCRENMIVLREGLEK 360
 301 IFHINYLGADQPTFLNFSKSSQVIAQYQLIRLGHGEGYRNVMENCRENMIVLREGLEK 360
 361 TERFNIVSKDEGVPLVAFSLKDSCHTEFEISDLRRYGWIVPAYTPPNAQHIITVLRVV 420
 361 TERFNIVSKENGVPVAVAFSLKDSCHTEFEISDLRRYGWIVPAYTPPNAQHIITVLRVV 420
 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGQESKSENSDMLVTVKSDIDK 480
 421 IREDFSRITLAERLVIDIEKVMRELDLPSRVIHKISLGQESKSENSDMLVTVKSDIDK 479
 481 Q--RDIIITGKVKVADRKTSTGI 501
 471 QREVYAVWKKFVDRKKTNGV 491

RESULT 11
 ABB91694
 ID ABB91694 standard; Protein; 500 AA.
 AC ABB91694;
 XX XX
 DT 31-MAY-2002 (first entry)
 XX XX
 DE Herbicidally active polypeptide SEQ ID NO 905.
 XX XX
 KW Herbicidal; plant; agriculture; herbicide.
 XX XX
 OS Arabidopsis thaliana.
 XX XX

WO200210210-A2.

07-FEB-2002.

28-AUG-2001; 2001WO-EP09892.

28-AUG-2001; 2001WO-EP09892.

(FARB) BAYER AG.

Tietjen K, Weidler M;

WPI; 2002-269010/31.

Identifying plant target compounds for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms -

Claim 5; SEQ ID NO 905; 261pp + Sequence Listing; English.

The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides.

Sequence 500 AA;

Query Match 79.5%; Score 2080; DB 23; Length 500; Best Local Similarity 78.8%; Pred. No. 2.8e-190; Matches 395; Conservative 44; Mismatches 52; Indels 10; Gaps 1;

1 MVLSHAVSESDVSHSTFASRYVTSLPRFKWPNSTIPKEAAQIINDELMDCGNPRLN 60
1 MVLSTASKSDSDSIHSTFASRYVRSNISRFEIPKNSIPKEAAQIINDELFKFGNPRNL 60
61 ASFVTTWMEPCDKLIMSSINKNYVDMDEYPTVTELNQRCVNMIAHLFNAPLEAEATAV 120
61 ASFVTTWMEPCDKLIMSSINKNYVDMDEYPTVTELNQRCVNMIAHLFNAPLEAEATAV 120
121 VGTGSSSEAIMLAGLAFKRWQNRKARBGKVPDKPNIVTGANVQVCWEKFARYFEVELK 180
121 VGTGSSSEAIMLAGLAFKRWQNRKARBGKVPDKPNIVTGANVQVCWEKFARYFEVELK 180
181 VKLSEGYVMPDPOQAVDMVDENTTCVADILGSTLNGEFEDVKLLNDLLVEKNKGTGWDTP 240
181 VKLSEGYVMPDPOQAVDMVDENTTCVADILGSTLNGEFEDVKLLNDLLVEKNKGTGWDTP 240
241 IHVDAASGGFTAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGVWVWTKDLDPEL 300
241 IHVDAASGGFTAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGVWVWTKDLDPEL 300
301 IFHNYLGADQPTFTLNFSGKSSQVIAQYQLIRLGHGFRVNVNMCNRENMI VLRGLEK 360
301 IFHNYLGADQPTFTLNFSGKSSQVIAQYQLIRLGHGFRVNVNMCNRENMI VLRGLEK 360
361 TERENIVSKDQGVPLVAFSLKDSCHTEFELSMLRNYGIVTPAYTPMPNAOHTVLRVV 420
361 TERENIVSKDQGVPLVAFSLKDSCHTEFELSMLRNYGIVTPAYTPMPNAOHTVLRVV 420
421 IREDFSRRLAERLVADIEKVMRELDLPSRVIHKISLGOEKSESNLMMVTKKSIDK 480
421 IREDFSRRLAERLVADIEKVMRELDLPSRVIHKISLGOEKSESNLMMVTKKSIDK 480
481 QRDIIITGWKXEVADRKKTSGI 501
471 QREVTAYKXKFDVTKDNGV 491

RESULT 12

AAO15140

ID AAO15140 standard; Protein; 502 AA.

XX AAO15140;

XX 25-SEP-2002 (first entry)

XX Tomato GAD protein.

XX GAD; plant GABA production regulation; glutamic acid decarboxylase; plant stress; GABA; gamma-aminobutyric acid; stress resistance.

XX Lycopersicon esculentum.

XX WO200238736-A2.

XX 16-MAY-2002.

XX 07-NOV-2001; 2001WO-US47447.

XX 07-NOV-2000; 2000US-246367P.

XX (EMER-) EMERALD BIOAGRICULTURE CORP.

XX Kinnersley AM, Turano FJ;

XX WPI; 2002-490073/52.

XX N-PSDB; AAL43418.

XX Making transformed plants that selectively increase gamma-aminobutyric acid production, by incorporating a DNA construct with a polynucleotide encoding a plant glutamic acid decarboxylase enzyme into plant's genome

XX Claim 17; Page 59; 63pp; English.

XX The present invention relates to a method of producing a transformed plant that selectively increases production of gamma-aminobutyric acid (GABA) in response to a signal, by incorporating into the plant's genome a DNA construct with a non-constitutive promoter operably linked to a polynucleotide encoding a functional plant glutamic acid decarboxylase (GAD), to provide a transformed plant that expresses the GAD coding sequence in response to a signal. Plants of this type have an enhanced ability to tolerate environmental or other stresses. The present sequence is the tomato GAD protein.

XX Sequence 502 AA;

XX Query Match 77.0%; Score 2014; DB 23; Length 502;

XX Best Local Similarity 76.0%; Pred. No. 6e-184;

XX Matches 387; Conservative 53; Mismatches 55; Indels 14; Gaps 5;

QY 1 MVLSHAVSESDVSHSTFASRYVTSLPRFKWPNSTIPKEAAQIINDELMDCGNPRLN 59

Db 1 MVLSTASKSDSDSIHSTFASRYVRSNISRFEIPKNSIPKEAAQIINDELFKFGNPRLN 60

QY 60 LASFVTTWMEPCDKLIMSSINKNYVDMDEYPTVTELNQRCVNMIAHLFNAPLEAEATAV 119

Db 61 LASFVTTWMEPCDKLIMSSINKNYVDMDEYPTVTELNQRCVNMIAHLFNAPLEAEATAV 120

QY 120 VGTGSSSEAIMLAGLAFKRWQNRKARBGKVPDKPNIVTGANVQVCWEKFARYFEVELK 179

Db 121 VGTGSSSEAIMLAGLAFKRWQNRKARBGKVPDKPNIVTGANVQVCWEKFARYFEVELK 180

QY 180 EVKLSSEGYVMPDPOQAVDMVDENTTCVADILGSTLNGEFEDVKLLNDLLVEKNKGTGWDTP 239

Db 181 EVKLSSEGYVMPDPOQAVDMVDENTTCVADILGSTLNGEFEDVKLLNDLLVEKNKGTGWDTP 240

QY 240 PIHVDAAASGGFTAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGVWVWTKDLDPEE 299

Db 241 PIHVDAAASGGFTAPFLYDLEWDFRLPLVKSINVSCHKYGLVYAGIGVWVWTKDLDPEE 300

with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides.

Sequence	494 AA;
Query Match	75.3%; Score 1969; DB 23; Length 494;
Best Local Similarity	75.2%; Pred. No. 1.2e-179;
Matches	373; Conservative 56; Mismatches 61; Indels 6; Gaps 3;
1	MVLSHAYSESDVSHSTFASRYRVSRLPFKPNPENSIPKEAAYQIINDELMLDGNPRLNL 60
1	MVLSA-TNSDSEHLHSTFASRYRVAVPPFRKPDHCPMDAAAYQVINDELMLDGNPRLNL 59
61	ASFTVTWMEPECDKLIMSSINKNYVDMDEYPTTQLQNRVNMIAHLFNPALPEAETAVG 120
60	ASFTVTWMEPECDKLIMDSVKNKYVDMDEYPTTQLQNRVNMIAHLFNPALPEAETAVG 119
121	VGVGSEATMLAGLAFKSKQWKRKAEGKPYDKENIVTGANVQVCKEKFARYPEVELKE 180
120	CGVGSSEATMLAGLAFKSKQWKRKAEGKPYDKENIVTGANVQVCKEKFARYPEVELKE 179
181	VKLSGGYVMDPQAVDMVDENITCVADILGSLTNGEFEDVKLLMDLVEKMKETGWDTP 240
180	VKLSGGYVMDPAKAVEMVDENITCVAAILGSLTNGEFEDVKQLNDLAEKNAETGWETP 239
241	IHVDAAGGGIAPLPELPELPELPELPELPELPELPELPELPELPELPELPELPEL 300
240	IHVDAAGGGIAPLPELPELPELPELPELPELPELPELPELPELPELPELPELPEL 299
301	IFHINYLGAQDPTFLNFKSGSSQVIAQYQIIRLGHGEGYRNVMENCRNNMIVLREGLEK 360
300	VFHINYLGAQDPTFLNFKSGSSQIIAQQYQIFIRLGHGEGYKNIMENCMNARRLREGIEM 359
361	TERFNIVSKDEGYVAFSLKSSCHTEEEISDMRRRWGIVPAPVMPFPAOHITVLRVV 420
360	TGRFNIVSKDGIQVPLVAFSLKSSKHTVEIAESLRKFGWIIIPATMPADAOHIAVLRVV 419
421	IREDFSRFLAERVIDIEKVMRELDLPSRVIAIKISLGOEKSESNLMMVTVKKSIDIK 480
420	IREDFSRGLADRLITHIQVLKIEGLSPSRIAHAAA---AAAVSGDDEEVKVTAKMSL 475
481	QRDIITGCKKFAVDRK 496
476	E-DITKTKVKKLVEHKR 490
RESULT 15	
AAAG13764	
ID	AAAG13764 standard; Protein; 494 AA.
XX	AAAG13764;
XX	AAAG13764;
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 13378.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.

PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135829.
PR	25-MAY-1999;	99US-0136021.
PR	25-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	14-JUN-1999;	99US-0138847.
PR	16-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.

Search completed: October 22, 2003, 15:05:26
Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

on on: October 22, 2003, 14:57:03 ; search time 44 Seconds
(without alignments)
1097.197 Million cell updates/sec

title: US-10-006-852-2
effect score: 2615
sequence: 1 MVLGSHAVSESDVSHSTFAS.....DIITGKKFVADRKKTSGIC 502

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

searched: 283308 segs, 9616862 residues

total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : PIR 76 : *
1: Piri : *
2: Pir2 : *
3: Pir3 : *
4: Pir4 : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2250	86.0	500	A48767	glutamate decarbox
2	2199	84.1	496	T01962	glutamate decarbox
3	2135.5	81.7	493	H84431	probable glutamate
4	2130	81.5	494	H95683	hypothetical prote
5	2080	79.5	500	G84431	probable glutamate
6	2014	77.0	502	S56177	probable glutamate
7	1318	50.4	514	JC7915	glutamate decarbox
8	1318	50.4	520	T49478	probable glutamate
9	1251.5	47.9	467	S75150	glutamate decarbox
10	1110	42.4	460	F70975	probable glutamate
11	1108	42.4	475	T36342	probable glutamate
12	1028	39.3	466	B85786	glutamate decarbox
13	1013	38.7	466	B43332	glutamate decarbox
14	1013	38.7	466	S24234	glutamate decarbox
15	1013	38.7	466	B90891	glutamate decarbox
16	1013	38.7	466	G81726	glutamate decarbox
17	1009	38.6	466	E91178	glutamate decarbox
18	1009	38.6	466	F86024	glutamate decarbox
19	983	36.8	467	AC1748	glutamate decarbox
20	955.5	36.5	462	AH1130	glutamate decarbox
21	953.5	36.5	467	AB1379	glutamate decarbox
22	948	36.3	464	AB1740	glutamate decarbox
23	940	35.9	464	AC1370	glutamate decarbox
24	887.5	33.9	585	S53072	glutamate decarbox
25	635.5	24.3	504	AE3623	glutamate decarbox
26	335.5	12.8	167	AF3623	glutamate decarbox
27	325	12.4	589	S70123	sphingosine-1-phos
28	314	12.0	384	G75102	group II decarboxy
29	312	11.9	363	E69015	glutamate decarbox

30	311	11.9	367	2	C69500	group II decarboxy
31	305	11.7	383	2	D71084	hypothetical prote
32	288	11.0	552	2	T33760	sphingosine-1-phos
33	270	10.3	544	2	C86405	probable sphingos
34	269.5	10.3	488	2	B69415	group II decarboxy
35	264.5	10.1	542	2	T29835	sphingosine-1-phos
36	249.5	9.5	355	2	D84192	glutamate decarbox
37	246.5	9.4	568	2	JC5923	sphingosine-1-phos
38	243.5	9.3	473	2	G72753	hypothetical prote
39	225	8.6	396	2	B64306	hypothetical prote
40	206.5	7.9	414	2	G69536	group II decarboxy
41	203	7.8	454	2	G72452	probable glutamate
42	147	5.6	557	2	H82481	decarboxylase, gro
43	144.5	5.5	548	2	H82234	probable glutamate
44	142.5	5.4	482	2	E96500	probable histidine
45	137	5.2	413	1	S39554	histidine decarbox

ALIGNMENTS

RESULT 1

A48767

Glutamate decarboxylase (EC 4.1.1.15), calmodulin-binding - garden petunia

C/Species: Petunia x hybrida (garden petunia)

C/Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 22-Jun-1999

C/Accession: A48767

R/Baum, G.; Chen, Y.; Azazi, T.; Takatsuji, H.; Fromm, H.

J. Biol. Chem. 268, 19610-19617, 1993

A/Title: A plant glutamate decarboxylase containing a calmodulin binding domain. Cloning

A/Reference number: A48767; MUID:93374956; PMID:8366104

A/Accession: A48767

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-500 <BAU>

A/Cross-references: GB:L16797; NID:G294111; PIDN:AAA33709.1; PID:G294112

C/Superfamily: Escherichia coli glutamate decarboxylase

C/Keywords: calmodulin binding; carbon-carbon lyase; carboxy-lyase

Query Match 86.0%; Score 2250; DB 2; Length 500;

Best Local Similarity 84.9%; Pred. No. 1.5e-150;

Matches 427; Conservative 39; Mismatches 33; Indels 4; Gaps 2;

Qy	1	MVLGSHAVSESDVSHSTFASRYVTSLSLPRFKMSPENSIPKEAAAYQIINDELMLDGNPRLN	60
Db	1	MVLGSHAVSESDVSHSTFASRYVTSLSLPRFKMSPENSIPKEAAAYQIINDELMLDGNPRLN	60
Qy	61	ASFTVTWMEPECDKLIWSSINKNYVDWDEYPTVTELRQRCVNMIAHLEWAPLEAEATAVG	120
Db	61	ASFTVTWMEPECDKLIWSSINKNYVDWDEYPTVTELRQRCVNMIAHLEWAPLEAEATAVG	120
Qy	121	VGTVGSSEATMLAGLAFKRQWNRKAEGKVPDKPNIVTGANVQVCWEKFPARYFEVELKE	180
Db	121	VGTVGSSEATMLAGLAFKRQWNRKAEGKVPDKPNIVTGANVQVCWEKFPARYFEVELKE	180
Qy	181	VKLSEGYVNDPQQAVDMDENTICVADILGSTLNGEFEDVKLLNDLLVEKKNKTEGWDTP	240
Db	181	VKLSEGYVNDPQQAVDMDENTICVADILGSTLNGEFEDVKLLNDLLVEKKNKTEGWDTP	240
Qy	241	IHVDAASGGFIAPFLYPELWDFRLPVKSKINVSCHKYGLVYAGIGWVWRNKEDLPEEL	300
Db	241	IHVDAASGGFIAPFLYPELWDFRLPVKSKINVSCHKYGLVYAGIGWVWRNKEDLPEEL	300
Qy	301	IFHINVLGADQPTFTLNFSKGSQVIAQYQLIRLGHGEGYRNVMNCRENMIIVLREGLEK	360
Db	301	IFHINVLGADQPTFTLNFSKGSQVIAQYQLIRLGHGEGYRNVMNCRENMIIVLREGLEK	360
Qy	361	TERPNIVSKDEGVPLVAFSLKDSCHTEFEISMLRRYGWIIVPAYTPNPAQHIIVLRVY	420
Db	361	TERPNIVSKDEGVPLVAFSLKDSCHTEFEISMLRRYGWIIVPAYTPNPAQHIIVLRVY	420
Qy	421	IREDFSRTLAERLVLDIEKVMRELDLPSRVHKKISLGQKESNSDNLMLVTVKKSIDIK	480
Db	421	IREDFSRTLAERLVLDIEKVMRELDLPSRVHKKISLGQKESNSDNLMLVTVKKSIDIK	480

b 421 IREDFSRTLAERLVRDIEKVLHELDLTPARVNAKLAABEQAAANGSEVH--KKTIDSEV 477
 y 481 ORDIIITGKKFKVAD-RKKTSGIC 502
 b 478 QLEMITAWLKFVDEKTKKTRVC 500

RESULT 2
 01962
 Lutamate decarboxylase (EC 4.1.1.15) 2, calmodulin-binding - common tobacco
 ,Species: Nicotiana tabacum (Common tobacco)
 ,Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 22-Jun-1999
 ,Accession: T01962
 ,Yun, S.J.; Oh, S.H.
 ,Cells 8, 125-129, 1998
 ,Title: Cloning and characterization of a tobacco cDNA encoding calcium/calmodulin-depe
 ,Reference number: Z14473; MUID:98302498; PMID:9638642
 ,Accession: T01962
 ,Status: preliminary; translated from GB/EMBL/DBJ
 ,Molecule type: mRNA
 ,Residues: 1-496 <YUN>
 ,Cross-references: EMBL:AF020424; NID:g3252853; PIDN:AAC39483.1; PID:g3252854
 ,Genetics:
 ,Gene: GAD2

Function:
 Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
 Superfamily: Escherichia coli glutamate decarboxylase
 Keywords: calmodulin binding; carbon-carbon lyase; carboxy-lyase

Query Match 84.1%; Score 2199; DB 2; Length 496;
 Best Local Similarity 84.1%; Pred. No. 5.8e-147;
 Matches 423; Conservative 34; Mismatches 38; Indels 8; Gaps 3;

1 MVLHAYSSDVSVHSTFASRYVRSLSLPRFKPENSIPKEAAAYQIINDELMDGNPRLNL 60
 1 MVLKTSASDVSVHSTFASRYVRSLSLPRFKPENSIPKEAAAYQIINDELMDGNPRLNL 60
 61 ASFVITWMEPCDKLIMSSINKNYVDMEYPTTELQNRVNNIAHLFNAPLEAEATVAG 120
 61 ASFVITWMEPCDKLIMSSINKNYVDMEYPTTELQNRVNNIAHLFNAPLEAEATVAG 120
 121 VGTVGSSEAIMLAGLAFKRWQNKRAQKGLPDKPNIVTGANVQVCEKFKARYFEVLEKE 180
 121 VGTVGSSEAIMLAGLAFKRWQNKRAQKGLPDKPNIVTGANVQVCEKFKARYFEVLEKE 180
 181 VKLSEGYVMDPQAVDMVDENITCVADILGSLTNGEFEDVKLLNDLIVKRNKGTGWDTP 240
 181 VKLSDGYVMDPEKAVEMVDENITCVAAALGSLTNGEFEDVKRLNDLITERNKGTGWDTP 240
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 301 IFRINYLGDQPTFTLNFSGKSSQVIAQYQIIRLGHGEGYRNVMENCRNMIIVLREGLEK 360
 301 IFRINYLGDQPTFTLNFSGKSSQVIAQYQIIRLGHGEGYRNVMENCRNMIIVLREGLEK 360
 361 TERFNIVSKDEGVPLVAFSLKDSCHTBEFEISDMLRRYGWIIVPMTMPNACHITVLRV 420
 361 TGRFKIVSKENGVPVAFSLKDSCHTBEFEISDMLRRYGWIIVPMTMPNACHITVLRV 420
 421 IREDFSRTLAERLVRDIEKVMRDELPSRVTHKISLQGEKSESDNLMVTVKSDIDK 480
 421 IREDFSRTLAERLVRDIEKVMRDELPSRVTHKISLQGEKSESDNLMVTVKSDIDK 480
 481 ORDIIITGKKFKVAD-RKKTSGIC 502
 474 QLEBITTAWLKFVADKTKKTRNGVC 496

C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C;Accession: H84431
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: H84431
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-493 <STO>
 A;Cross-references: GB:AE002093; NID:g4406789; PIDN:AAD20099.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g02010
 A;Map position: 2
 C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 81.7%; Score 2135.5; DB 2; Length 493;
 Best Local Similarity 82.4%; Pred. No. 1.7e-142;
 Matches 411; Conservative 32; Mismatches 45; Indels 11; Gaps 2;

Qy 1 MVLHAYSSDVSVHSTFASRYVRSLSLPRFKPENSIPKEAAAYQIINDELMDGNPRLNL 60
 Db 1 MVLKTSASDVSVHSTFASRYVRSLSLPRFKPENSIPKEAAAYQIINDELMDGNPRLNL 60
 Qy 61 ASFVITWMEPCDKLIMSSINKNYVDMEYPTTELQNRVNNIAHLFNAPLEAEATVAG 120
 Db 61 ASFVITWMEPCDKLIMSSINKNYVDMEYPTTELQNRVNNIAHLFNAPLEAEATVAG 120
 Qy 121 VGTVGSSEAIMLAGLAFKRWQNKRAQKGLPDKPNIVTGANVQVCEKFKARYFEVLEKE 180
 Db 121 VGTVGSSEAIMLAGLAFKRWQNKRAQKGLPDKPNIVTGANVQVCEKFKARYFEVLEKE 180
 Qy 181 VKLSEGYVMDPQAVDMVDENITCVADILGSLTNGEFEDVKLLNDLIVKRNKGTGWDTP 240
 Db 181 VKLSDGYVMDPEKAVEMVDENITCVAAALGSLTNGEFEDVKRLNDLITERNKGTGWDTP 240
 Qy 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 Db 241 IHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIGWVWRNKEDLPEEL 300
 Qy 301 IFRINYLGDQPTFTLNFSGKSSQVIAQYQIIRLGHGEGYRNVMENCRNMIIVLREGLEK 360
 Db 301 IFRINYLGDQPTFTLNFSGKSSQVIAQYQIIRLGHGEGYRNVMENCRNMIIVLREGLEK 360
 Qy 361 TERFNIVSKDEGVPLVAFSLKDSCHTBEFEISDMLRRYGWIIVPMTMPNACHITVLRV 420
 Db 361 TGRFKIVSKENGVPVAFSLKDSCHTBEFEISDMLRRYGWIIVPMTMPNACHITVLRV 420
 Qy 421 IREDFSRTLAERLVRDIEKVMRDELPSRVTHKISLQGEKSESDNLMVTVKSDIDK 480
 Db 421 IREDFSRTLAERLVRDIEKVMRDELPSRVTHKISLQGEKSESDNLMVTVKSDIDK 480
 Qy 481 ORDIIITGKKFKVAD-RKKTSGIC 499
 Db 471 QREVTAYWKKLL-ETKKTIN 488

RESULT 4
 H96683
 Hypothetical protein F12P19_12 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C;Accession: H96683
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

SULT 3
 4431
 obable glutamate decarboxylase [imported] - Arabidopsis thaliana

120 GYGTGSSBAIMLAGLAFKRWKQKKAEGKPKVDKPNIVTGANVQCWEKFAFYFEVELK 179
 121 GYGTGSSBAIMLAGLAFKRWKQKKAEGKPKVDKPNIVTGANVQCWEKFAFYFEVELK 180
 180 EVKLSGYYVMPDQQAVDMDVENTICVADILGSLNGEFEDVKLLNDLLEKKNKFTGMDT 239
 181 EVKLSGYYVMPDPAKAVEIVDENTICVAAIILGSLTNGEFEDVKLLNELLTKKNKFTGMDT 240
 240 PIHVDAASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIYGVINWENKEDLPER 299
 241 PIHVDAASGGFIAPFLWPLEWDFRPLVKSINVSCHKYGLVYAGIYGVINWENKEDLPER 300
 300 LIPHINYLGAQDQFTFLNFKGSSQVIAQYQILRLHGHEGYRNVNMCNENMIVLREGLE 359
 301 LVPHINYLGSQDQFTFLNFKGSSQVIAQYQILRLHGHEGYRNVNMCNENMIVLREGLE 360
 360 KTERFNIVSKDEGVPLVAFSLKSSCHTEFEISDMMLRRYGVIPVAYTTPPNAQHTVLRV 419
 361 KMGRFDIVSKDVGVPVYAFSLRDSKSYTFVSEBHLRFRFGWIVPAYTTPPDAEHTAVLRV 420
 420 VIREDFSRTLAERLVIDIEKVMREDELSPRVTHKI--SLGQKSESNDNL---WVTV 473
 421 VIREDFSRSLAERLVIDIEKILSELDTQPRPTKAVRVTAEBVRDDKGDGLHFFHMDTV 480
 474 KKSDDIKQRDIITGKWKFFVADRKKTSGIC 502
 481 -----ETOKDIIKHWRKIAG--KKTSGVC 502

RESULT 7

JC7915
 glutamate decarboxylase (EC 4.1.1.15) - Aspergillus oryzae
 Species: Aspergillus oryzae
 Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 21-Apr-2003
 Accession: JC7915
 Kato, Y., Kato, Y., Furukawa, K., and Hara, S.
 J. Biol. Chem. 276: 2600-2605, 2001
 Title: Cloning and nucleotide sequence of the glutamate decarboxylase-encoding gene ga
 , Reference number: JC7915; PMID:22484577; PMID:12596854
 Accession: JC7915
 Molecule type: DNA
 Residues: 1-514 <KAT>
 Cross-references: DDBJ:AB025422
 Comment: This enzyme produces gamma-aminobutyric acid.
 Gene: gada
 Introns: 79/2; 88/3; 477/2
 Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 50.4%; Score 1318; DB 2; Length 514;
 Best Local Similarity 56.6%; Pred. No. 6e-85;
 Matches 256; Conservative 77; Mismatches 109; Indels 10; Gaps 5;
 / 8 SESDVSVHSTFASRYVTSLPREFKMPENSIKPEAAAYOIINDELMLDGNLRLNLAGSVTTW 67
 / 28 TDEDEGYSYVGRFATEQLPQEMEREMPREVAVRMKDELSDGNLRLNLAGSVTTW 87
 / 68 MEPECDKLIIMSSINKNVDMDEYPTTELQNRVNVNIAHLFNAPLEAE--EABTAVGVTVGS 126
 / 88 MEDEAEKLMTESEPKNFIDYEEYPOQAEIQRNCRVNVNIAHLFNAPVHSEDEHPMGTSTIGS 147
 / 127 SEAIMLAGLAFKRWKQKKAEGKPKVDKPNIVTGANVQCWEKFAFYFEVELKLSBG 186
 / 148 SEAIMLGLVLAAMKRWKQKKAEGKPKVDKPNIVMNSAVQCWEKFAFYFEVEKFTVCTP 207
 / 187 YYVMDPQQAVDMDVENTICVADILGSLNGEFEDVKLLNDLLEKKNKFTGMDTPIHVDA 246
 / 208 RYVIDPQQAVDMDVENTIGICAILGTYTTEYEDVKAINDLIERN---IDVPIHVDA 263
 / 247 SGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIYGVINWENKEDLPELPHIN 306
 / 264 SGGFVAPFIPKLEWDFRPLVKSINVSCHKYGLVYAGIYGVINWENKEDLPELPHIN 323

Qy 307 LGADQPTFLNFKGSSQVIAQYQILRLHGHEGYRNVNMCNENMIVLREGLEKTERPNI 366
 Db 324 LGAEQASFTLNFKGASQVIGQYQIRLQKGRYSIMINIVTADFLAQELEKM-GFTI 382
 Qy 367 VS--KDEGVPLVAFSL--KSSCHTEFEISDMMLRRYGVIPVAYTTPPNAQHTVLRV 422
 Db 383 MSQRGHGLPLVAFRLPAEREGQDFALAHQHLRERGVIPVAYTMAPNSNKLKLMRVVVR 442
 Qy 423 EDFSRTLAEELVIDIEKVMREDELSPRVTHKI 454
 Db 443 EDFKSRCDALLSDIKLGLKLTGMDKAMDK 474

RESULT 8

T49478
 Probable glutamate decarboxylase [imported] - Neurospora crassa
 N; Alternate names: protein B14D6.300
 C; Species: Neurospora crassa
 C; Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
 C; Accession: T49478
 R; Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, May 2000
 A; Reference number: Z25022
 A; Accession: T49478
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-520 <SCH>
 A; Cross-references: EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.300
 A; Experimental source: BAC clone B14D6; strain OR74A
 C; Genetics:
 A; Gene: NCSP:B14D6.300
 A; Map position: 6
 A; Introns: 85/2; 504/3
 C; Superfamily: Escherichia coli glutamate decarboxylase

Query Match 50.4%; Score 1318; DB 2; Length 520;
 Best Local Similarity 55.6%; Pred. No. 6.1e-85;
 Matches 262; Conservative 73; Mismatches 118; Indels 18; Gaps 6;
 Qy 8 SESDVSVHSTFASRYVTSLPREFKMPENSIKPEAAAYOIINDELMLDGNLRLNLAGSVTTW 67
 Db 34 NDEDRFTSYGSKFAAADLPRHEMPEDEMPKAVRMKDELSDGNLRLNLAGSVTTW 93
 Qy 68 MEPECDKLIIMSSINKNVDMDEYPTTELQNRVNVNIAHLFNAPLEAE--TAVGVTVG 125
 Db 94 MEEBAEKLMTESEPKNFIDYEEYPOQAEIQRNCRVNVNIAHLFNAPVHSEDEHPMGTSTIGS 153
 Qy 126 SSEAIMLAGLAFKRWKQKKAEGKPKVDKPNIVTGANVQCWEKFAFYFEVELKLS 185
 Db 154 SSEAIMLGLVLAAMKRWKQKKAEGKPKVDKPNIVMNSAVQCWEKFAFYFEVEKFTVCTP 213
 Qy 186 GYYVMDPQQAVDMDVENTICVADILGSLNGEFEDVKLLNDLLEKKNKFTGMDTPIHVDA 245
 Db 214 DRYVIDPQAVDMDVENTIGICAILGTYTTEYEDVKAINDLIERN---GLDTPIHVDA 269
 Qy 246 ASGGFIAPFLYPELEWDFRPLVKSINVSCHKYGLVYAGIYGVINWENKEDLPELPHIN 305
 Db 270 ASGGFVAPFIPKLEWDFRPLVKSINVSCHKYGLVYAGIYGVINWENKEDLPELPHIN 329
 Qy 306 YLGADQPTFLNFKGSSQVIAQYQILRLHGHEGYRNVNMCNENMIVLREGLEKTERFN 365
 Db 330 YLGADQASFTLNFKGASQVIGQYQIRLQKGRYSIMINIVTADFLAQELEKM-GFTI 388
 Qy 366 IVSK--DEGVPLVAFSLKSSCHTE--EFEISDMMLRRYGVIPVAYTTPPNAQHTVLRV 421
 Db 389 IMSQKSGQLPLVAFRLKEDPDRYDFALAHQHLRERGVIPVAYTMAPKTEGLKMLRIVV 448
 Qy 422 REDESRTLAERLVIDI-----EKVMREDELSPRVTHKISLGQKSESNDNL 465
 Db 449 REDFSRNRCDGLISDIRSCQGLEQMDKETVKKQOEFHKKHVVVSGKASHN 499

b 40 PAHRLPDSPLPPESAYRLVHDELMLDGNARLNLAFTVTTWMEPQAGVLMSECRDKNMIDK 99
 y 88 DEYPTTELOQRVNMIAHLFNAPLEAEATAVGVGVSSEAIMLAGLAFKRW--QNKR 145
 b 100 DEYPTAELEBRCCVAMDLWHP--DPSTAVGCSPTGSSSEACMLAGWALRRWRARNAD 157
 y 146 KAREKPKVDNITGANVQVCKEFARYFEVELKEVYKLSGYYVMDPQAAVMDVDENTIC 205
 b 158 RYPAKV--RENLVMGVNVQVCKFKCFNFEVEARQVPMGDRFHLDPQAAAELCDENIG 216
 y 206 VADILGSLNGEPEFEDVKNLNDLIVKKNKTKTGMWTPPHVDAASGGFIAPFLYPELSDWDR 265
 b 217 VVGLGSLGTFDGSYEPVADLCAALDALQERTGLDVPVHVVDGASGAMVAPPELDDLVMDR 276
 y 266 PLVKSINVSCHKYGLVYVGVWALNRKEDLPBELIFHINVLGADOPTFLNFSKGSQV 325
 b 277 PRVASINTSCHKYGLVYVGVWALNRKEDLPBELIFHINVLGADOPTFLNFSRFGQV 336
 y 326 IAQYQILRILGHEGYRNVNENCRNMIIVREGLEKTERENIVSKDEGVPLVAFSLKDS-S 384
 b 337 VAQYVNFRLREGYRVAQOSARDIAGSLAERVAALGDFRLLTRGDLPVFAFTTADDT 396
 y 385 CHTEFEISDMLRYGVIPAYTTPNNAOHTVLRVIREDFSRITLAERLVIDIEKVMREL 444
 c 397 AYDVFDSRRLREGGWLVPAYTTPPHRELDLVLURVCRNGFSADWADLLADLERLLPEL 456
 y 445 DELP 448
 b 457 RROP 460

RESULT 12
 86786
 lutamate decarboxylase (EC 4.1.1.15) [imported] - Lactococcus lactis subsp. lactis (str
 ,Species: Lactococcus lactis subsp. lactis
 ,Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 ,Accession: S86786
 ,Boloitin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli
 , genome Res. 11, 731-753, 2001
 ,Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 ,Reference number: A86625; MUID:21235186; PMID:11337471
 ,Accession: S86786
 ,Status: preliminary
 Molecule type: DNA
 Residues: 1-466 <SMT>
 Cross-references: GB:AE005176; PID:g12724267; PIDN:AAK05388.1; GSPDB:GN00146
 Experimental source: strain IL1403
 Genes:
 Superfamily: Escherichia coli glutamate decarboxylase
 Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 39.3%; Score 1028; DB 2; Length 466;
 Best Local Similarity 45.0%; Pred. No. 1.3e-64;
 Matches 195; Conservative 91; Mismatches 141; Indels 6; Gaps 3;
 / 18 FASRYVTSLPRFKPENSIPKEAAYQIINDELMDGNPRNLINLASFVTTWMEPCDKLIM 77
 c 18 FGSSEQVDFPKYKLAQCSIEPRVAVQVQVDEMLDGNARLNLAFTCYMPEAVKLM 77
 / 78 SSINKNVYDMDYEVVTTTELQNRVNMIAHLFNAPLEAEATAVGVGVSSEAIMLAGLAF 137
 c 78 QTLEKNALDKSEYPTTEIENRCVNMIAHLWNA--SEKXFMGTSTIGSSEACMLGGMAM 135
 y 138 KRKQWKRKAEKPKVD--KNIVTGANVQVCKEFARYFEVELKEVYKLSGYYVMDPQAA 195
 c 136 KFSWRKRAEKLGLDINAKPNLNTSSGQVCKEFARYFEVELKEVYKLSGYYVMDPQAA 195
 y 196 VMDYDVENTICVADILGSLNGEPEFEDVKNLNDLIVKKNKTKTGMWTPPHVDAASGGFIAP 255
 c 196 MDYVDEYTIYGVVGLMGLTYTGRYDIDIKALDNLLEBYNKNQTDYKVIHVDAAASGLYAPV 255
 y 256 YPELWDFRPLVKSINVSCHKYGLVYVGVWALNRKEDLPBELIFHINVLGADOPTFT 315

Db 256 EPELWDFRPLKWNISINTSCHKYGLVYVGVWALNRKEDLPBELIFKYSYLGELPTWA 315
 Qy 316 LNFSSKSSQVIAQYQIILRILGHEGYRNVNENCRNMIIVREGLEKTERENIVSKDEGVPL 375
 Db 316 INFSGASQLIGQYVNFYRFGDYKAIHERTHKVAWFLAKEIEKTKGMEFIMWDSQLPI 375
 Qy 376 VAFSLKDSCH--TEFEISDMLRYGVIPAYTTPNNAOHTVLRVIREDFSRITLAERL 433
 Db 376 VCYYKLEDSNRGNWLYDLADRLMLKMGQVPAVPLPKNLENIQRLVIRADFGMNAFNY 435
 Qy 434 VIDIEKVMREDE 446
 Db 436 VDMQEAIEALNK 448

RESULT 13
 843332
 glutamate decarboxylase (EC 4.1.1.15) beta - Escherichia coli (strain K-12)
 N:Alternate names: L-glutamate 1-carboxy-lyase
 C:Species: Escherichia coli
 C:Date: 10-Mar-1994 #sequence_revision 23-Mar-1995 #text_change 01-Mar-2002
 C:Accession: B43332; S30261; H64902
 R:Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
 J. Bacteriol. 174, 5820-5826, 1992
 A:Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to
 A:Reference number: A43332; MUID:92394884; PMID:1522060
 A:Accession: B43332
 A:Molecule type: DNA
 A:Residues: 1-466 <SMT>
 A:Cross-references: GS:M84025; NID:g146059; PIDN:AAA23834.1; PID:g146060
 R:Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
 Mol. Gen. Genet. 237, 113-122, 1993
 A:Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis of
 A:Reference number: S30261; MUID:93204884; PMID:8455549
 A:Accession: S30261
 A:Molecule type: protein
 A:Residues: 1-4, LQVL, /-15 <YOS>
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Ferna, N.T.; Burland, V.; Riley, M.; Cc
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278903
 A:Accession: H64902
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-466 <BLAT>
 A:Cross-references: GB:AE000246; GB:U00096; NID:g1787764; PIDN:AACT4566.1; PID:g1787769,
 A:Experimental source: strain K-12, substrain MGI655
 C:Genetics:
 A:Gene: gadB
 A:Map position: 78 min
 C:Function:
 A:Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
 A>Note: in E. coli, two isoforms (alpha and beta) have been found, each encoded by a sep
 C:Superfamily: Escherichia coli glutamate decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal i
 F:2/76/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 38.7%; Score 1013; DB 1; Length 466;
 Best Local Similarity 45.0%; Pred. No. 1.5e-63;
 Matches 197; Conservative 83; Mismatches 152; Indels 6; Gaps 3;
 Qy 16 STFASRYVTSLPRFKPENSIPKEAAYQIINDELMDGNPRNLINLASFVTTWMEPCDKL 75
 Db 16 SRFGAKSITIAESKRFPLEHMRDDVAQIINDELMDGNARLNLAFTCYMPEAVKLM 75
 Qy 76 IMSSINKNVYDMDYEVVTTTELQNRVNMIAHLFNAPLEAEATAVGVGVSSEAIMLAGL 135
 Db 76 MDLSEKNALDKSEYPTTEIENRCVNMIAHLWNA--SEKXFMGTSTIGSSEACMLGGM 135
 Qy 136 AFKRKQWKRKAEKPKVD--KNIVTGANVQVCKEFARYFEVELKEVYKLSGYYVMDPQAA 195


```

Matches 137; Conservative 83; Mismatches 152; Indels 6; Gaps 3;
Y 16 STFASRYVTSLSRFRKMPENSIPKAAAYQIINDELMLDGNPRINLASFVTTWMEPECDKL 75
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 16 SRFGASISITIAESKRFPLEHMRDDVAFQIINDELYLDGNARQNLATFCQIWDDENVHKL 75
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 76 IMSINKNVYVMDYPTTELQNRVYVMIHLFNAFLPEAEATAVGVTVGSSEAIMLAGL 135
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 76 MDLSINKNWKDEEYQSAADLRCVNVVADLWHPAPKNGQAVGNTIGSSEACMLGGM 135
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 136 AFKKQNKKAEGKVDYDNIVTGANVQVCWEKFAFYFELKEVKLSEGYVVDPOQA 195
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 136 AMKWRWRKREAAAGKFTDKENLVCGP-VQICWHKFARYWDVLRBIPMRPGQLFMDPKRM 194
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 196 VDMVDENTICVADILGSTLNGFEDVKLNDLLVEKNKETGWDTPIHVDAASCGFIAPPL 255
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 195 IEACDENTIGVPTFGVYTCNVYFPPLHDALDKFQADTGIDIDMHDIDAASGGFLAPFV 254
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 256 YPELEWDFRLPLVKSINVSCHKYGLVYAGIYVWVWRNKEDLPEELIFHINYLGADQPTFT 315
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 255 APDIWDFRLPRVKSISASGHKFLAPLGGGWVWFRDEBALPQELVFNVDYLLGGQIGTFA 314
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 316 LNFSKGSQVIAQYYQLIRLGHGGRVNMENCRNMIVLREGLEKTERENIV---SKDQG 372
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 315 INFSPAGQVIAQYIEFLRLREGYTKVQNASYQVAAYLADEIAKLGPYEFFCTGRPDEG 374
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 373 VPLVAFSLKDS--CHTEFEISDMLRFRYVVPAYTMPENAQHITVLRVVIREFDSRTLA 430
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 375 IPAVCFKLXGDGPPGYTLYDLSERLRLRGVQVPAFTLGGSEATDIVVMRIMCRGFEMDFA 434
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 431 ERLVIDIEKVMRELELP 448
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Y 435 ELLEDYKASLKYLSDHP 452
b : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: October 22, 2003, 15:08:36
 00 time : 45 secs